Contents

Preface **16**

1 Chapter 1 **18**

Connectives and Derivations **18**
1.0.1 Roadmap . . . 19
1.0.2 Named moves . . . 19
1.0.3 Summary . . . 20
1.0.4 Terms . . . 20

1.1 Statements, Arguments and How to Break One **20**
1.1.1 Two arguments about squares . . . 20
1.1.2 Statements . . . 21
1.1.3 The connectives . . . 23
1.1.4 When is an implication broken? . . . 24
1.1.5 "Only if", "necessary" and "sufficient" . . . 24
1.1.6 Translating a sentence . . . 25
1.1.7 A phrase book . . . 25
1.1.8 Reading a statement: precedence and the main connective . . . 27
1.1.9 The form of an argument . . . 28
1.1.10 Try to break it . . . 29
1.1.11 Good form, false sentences . . . 30
1.1.12 Two ways to attack an argument . . . 31
1.1.13 Each connective, two rules . . . 32
1.1.14 What comes next . . . 32
1.1.15 Exercises . . . 33
1.1.15.1 A. Check your understanding . . . 33
1.1.15.2 B. Practice . . . 35
1.1.15.3 C. Going deeper . . . 39

1.2 Derivations, Starting with And **40**
1.2.1 Why check by matching . . . 40
1.2.2 A rule is a pattern . . . 41
1.2.3 Derivations without boxes . . . 42
1.2.4 First derivations . . . 43
1.2.5 Reading a derivation . . . 44
1.2.6 What a derivation proves . . . 45
1.2.7 The rules for and . . . 47
1.2.8 Two directions of reasoning . . . 49
1.2.9 First derivations with and . . . 49
1.2.9.1 A line is not used up . . . 52

1.2.10 How this looks in a mathematical proof . . . 54
1.2.11 Common errors . . . 55
1.2.12 The rule card . . . 56
1.2.13 Exercises . . . 57
1.2.13.1 A. Check your understanding . . . 57
1.2.13.2 B. Practice . . . 59
1.2.13.3 C. Going deeper . . . 61

1.3 Using an Implication: Modus Ponens 65
1.3.1 Four cards . . . 65
1.3.2 The rule . . . 66
1.3.3 Chains of implications . . . 69
1.3.4 Checking a line of modus ponens . . . 73
1.3.5 Two chains about numbers . . . 73
1.3.6 How this looks in a mathematical proof . . . 77
1.3.7 Common errors . . . 77
1.3.8 The rule card . . . 78
1.3.9 Exercises . . . 79
1.3.9.1 A. Check your understanding . . . 79
1.3.9.2 B. Practice . . . 81
1.3.9.3 C. Going deeper . . . 85

1.4 Proving an Implication: the Rule of the Box 87
1.4.1 Supposing something we do not know . . . 87
1.4.2 The introduction rule for implication . . . 88
1.4.3 A first box: implication is transitive . . . 89
1.4.4 Argument B, diagnosed . . . 91
1.4.5 A box must end on a line of its own . . . 93
1.4.6 Which lines a line may cite . . . 93
1.4.7 How to check a derivation . . . 96
1.4.8 An assumption used outside its box . . . 96
1.4.9 More boxes: from nothing, and one inside another . . . 98
1.4.10 Two hypotheses, or one at a time . . . 99
1.4.11 The rule card . . . 102
1.4.12 Exercises . . . 104
1.4.12.1 A. Check your understanding . . . 104
1.4.12.2 B. Practice . . . 106
1.4.12.3 C. Going deeper . . . 109

1.5 Or, Introduced 115
1.5.1 Throwing information away on purpose . . . 115
1.5.2 The rule of or-introduction . . . 116
1.5.3 A disjunction as a goal . . . 117
1.5.4 Manufacturing a hypothesis . . . 122
1.5.5 A goal we cannot reach yet . . . 124
1.5.6 Derive it or break it . . . 124
1.5.7 How this looks in a mathematical proof . . . 125
1.5.8 The rule card . . . 126
1.5.9 Exercises . . . 127
1.5.9.1 A. Check your understanding . . . 127
1.5.9.2 B. Practice . . . 129
1.5.9.3 C. Going deeper . . . 132

1.6 Using Or: Proof by Cases **134**
1.6.1 A proof that needs two cases . . . 134
1.6.2 The rule of or-elimination . . . 135
1.6.3 Cases end in the current goal . . . 136
1.6.4 Two cases, two conclusions . . . 140
1.6.5 Laws of "or" . . . 142
1.6.6 From a skeleton to a proof about numbers . . . 146
1.6.7 The rules of this chapter . . . 150
1.6.7.1 Named moves . . . 153
1.6.8 Exercises . . . 154
1.6.8.1 A. Check your understanding . . . 154
1.6.8.2 B. Practice . . . 157
1.6.8.3 C. Going deeper . . . 169
1.6.8.4 R. Review . . . 174

2 Chapter 2 **177**

Negation and Classical Logic **177**
2.0.1 Roadmap . . . 177
2.0.2 Named moves . . . 178
2.0.3 Summary . . . 178
2.0.4 Terms . . . 178

2.1 Contradiction: Using a Negation **179**
2.1.1 A case that cannot occur . . . 179
2.1.2 The rules for a clash . . . 179
2.1.3 First derivations with a clash . . . 182
2.1.4 How the rules look in a mathematical proof . . . 187
2.1.5 Vacuous truth . . . 189
2.1.6 Two sequents that sound wrong . . . 190
2.1.7 The rule card . . . 191
2.1.8 Exercises . . . 192
2.1.8.1 A. Check your understanding . . . 192
2.1.8.2 B. Practice . . . 193
2.1.8.3 C. Going deeper . . . 199

2.2 Proving a Negation **201**
2.2.1 How to prove that something is not so . . . 201
2.2.2 The first example: a number that is never zero . . . 203
2.2.3 First derivations with not-introduction . . . 204
2.2.4 How the rule looks in a mathematical proof . . . 207
2.2.5 The rule card . . . 209
2.2.6 Exercises . . . 210
2.2.6.1 A. Check your understanding . . . 210
2.2.6.2 B. Practice . . . 212
2.2.6.3 C. Going deeper . . . 218

2.3 Derived Rules: Proved Sequents as New Rules **219**
2.3.1 A box that keeps coming back . . . 219
2.3.2 Modus tollens, from the rules we have . . . 220
2.3.3 What a derived rule is . . . 222

2.3.4 Pasting whole derivations: weakening and cut . . . 224
2.3.5 Disjunctive syllogism: the impossible case, filed . . . 226
2.3.6 Modus tollens on numbers, and the green card . . . 229
2.3.7 How this rule looks in a mathematical proof . . . 230
2.3.8 Common errors . . . 231
2.3.9 The rule card . . . 232
2.3.10 Exercises . . . 233
2.3.10.1 A. Check your understanding . . . 233
2.3.10.2 B. Practice . . . 235
2.3.10.3 C. Going deeper . . . 240

2.4 Iff, Equivalence, and What a Definition Says 243
2.4.1 The rules for "if and only if" . . . 244
2.4.2 First derivations with iff . . . 245
2.4.3 Provable equivalence . . . 251
2.4.4 Proving an iff: two directions . . . 253
2.4.5 What a definition says . . . 255
2.4.6 Chains of equivalences . . . 257
2.4.7 The rule card . . . 259
2.4.8 Exercises . . . 261
2.4.8.1 A. Check your understanding . . . 261
2.4.8.2 B. Practice . . . 262
2.4.8.3 C. Going deeper . . . 269

2.5 Proof by Contradiction and the Classical Rules 270
2.5.1 Two boxes that look alike . . . 270
2.5.2 Where the rules so far stop short . . . 271
2.5.3 Reductio ad absurdum . . . 271
2.5.4 The move: when stuck, assume the negation of the goal . . . 273
2.5.5 Derivations with the new rule . . . 273
2.5.6 De Morgan's laws . . . 279
2.5.7 Implication as a disjunction . . . 284
2.5.8 One principle in three forms . . . 286
2.5.9 How reductio looks in a mathematical proof . . . 290
2.5.10 The rule card . . . 291
2.5.11 Exercises . . . 293
2.5.11.1 A. Check your understanding . . . 293
2.5.11.2 B. Practice . . . 296
2.5.11.3 C. Going deeper . . . 304
2.5.11.4 R. Review . . . 309

3 Chapter 3 313

Finding Derivations 313
3.0.1 Roadmap . . . 313
3.0.2 Named moves . . . 314
3.0.3 Summary . . . 314
3.0.4 Terms . . . 314

3.1 How to Find a Derivation 315
3.1.1 The moves so far . . . 315

3.1.2 Partial derivations . . . 316
3.1.3 The method . . . 317
3.1.4 Splitting the goal . . . 320
3.1.5 Cases must end in the goal . . . 322
3.1.6 Postponing a disjunction . . . 324
3.1.7 When nothing else works . . . 326
3.1.8 Four refinements . . . 329
3.1.9 A search that fails . . . 331
3.1.10 A mathematical goal . . . 332
3.1.11 The rules in one place . . . 334
3.1.12 Exercises . . . 338
3.1.12.1 A. Check your understanding . . . 338
3.1.12.2 B. Practice . . . 340
3.1.12.3 C. Going deeper . . . 344

3.2 Four Kinds of Task 349
3.2.1 Four tasks, one method . . . 350
3.2.2 Theorems . . . 350
3.2.3 Equivalences . . . 351
3.2.4 Inconsistent lists . . . 353
3.2.5 The counterexample set, derived . . . 357
3.2.6 One task in four costumes . . . 358
3.2.7 Derive or break . . . 359
3.2.8 Exercises . . . 361
3.2.8.1 A. Check your understanding . . . 361
3.2.8.2 B. Practice . . . 363
3.2.8.3 C. Going deeper . . . 369

3.3 From Argument to Derivation 371
3.3.1 One key for the section . . . 371
3.3.2 Extracting an argument . . . 372
3.3.3 Two columns . . . 375
3.3.4 A library of standard sequents . . . 377
3.3.5 Exercises . . . 379
3.3.5.1 A. Check your understanding . . . 379
3.3.5.2 B. Practice . . . 379
3.3.5.3 C. Going deeper . . . 387
3.3.5.4 R. Review . . . 390

4 Chapter 4 394

Truth, Soundness and Completeness 394
4.0.1 Roadmap . . . 395
4.0.2 Named moves . . . 395
4.0.3 Summary . . . 396
4.0.4 Terms . . . 396

4.1 Formulas as Objects 397
4.1.1 Why formulas must become objects . . . 397
4.1.2 The alphabet . . . 398
4.1.3 Formulas . . . 399

4.1.4 Why this definition . . . 403
4.1.5 Parse trees . . . 404
4.1.6 Subformulas . . . 405
4.1.7 Shorthand and official formulas . . . 406
4.1.8 Letters, metavariables and substitution . . . 407
4.1.9 Exercises . . . 409
4.1.9.1 A. Check your understanding . . . 409
4.1.9.2 B. Practice . . . 411
4.1.9.3 C. Going deeper . . . 412

4.2 Induction on the Natural Numbers 413
4.2.1 One argument for infinitely many cases . . . 413
4.2.2 The principle of induction . . . 413
4.2.3 Strong induction . . . 417
4.2.4 Proving more to prove less . . . 418
4.2.5 Exercises . . . 419
4.2.5.1 A. Check your understanding . . . 419
4.2.5.2 B. Practice . . . 421
4.2.5.3 C. Going deeper . . . 423

4.3 Induction and Recursion on Formulas 424
4.3.1 Structural induction . . . 424
4.3.2 Counting parentheses: the depth profile . . . 427
4.3.3 Unique readability . . . 430
4.3.4 Definition by recursion . . . 432
4.3.5 Exercises . . . 435
4.3.5.1 A. Check your understanding . . . 435
4.3.5.2 B. Practice . . . 437
4.3.5.3 C. Going deeper . . . 439

4.4 Valuations and Truth Tables 444
4.4.1 Reading a table off the rules . . . 445
4.4.2 Valuations . . . 446
4.4.3 The value of a formula . . . 446
4.4.4 Only the atoms of a formula matter . . . 449
4.4.5 Truth tables . . . 451
4.4.6 Truth-functional connectives . . . 453
4.4.7 Exercises . . . 455
4.4.7.1 A. Check your understanding . . . 455
4.4.7.2 B. Practice . . . 457
4.4.7.3 C. Going deeper . . . 460

4.5 Tautologies, Entailment and Equivalence 463
4.5.1 Tautologies and contradictions . . . 464
4.5.2 Semantic entailment . . . 465
4.5.3 Semantic equivalence . . . 469
4.5.4 Relations between two formulas . . . 472
4.5.5 Replacing a part by an equivalent part . . . 473
4.5.6 Substituting formulas for atoms . . . 475
4.5.7 One row or every row . . . 477
4.5.8 Two questions, side by side . . . 477
4.5.9 Exercises . . . 478

4.5.9.1 A. Check your understanding . . . 478
4.5.9.2 B. Practice . . . 479
4.5.9.3 C. Going deeper . . . 482

4.6 Hunting for a Refuting Row 485
4.6.1 The rules of the search . . . 486
4.6.2 The search, drawn as a tree . . . 486
4.6.3 What a closed tree proves . . . 489
4.6.4 Choosing the next move . . . 491
4.6.5 Testing a list for satisfiability . . . 492
4.6.6 The withheld verdicts . . . 494
4.6.7 Exercises . . . 495
4.6.7.1 A. Check your understanding . . . 495
4.6.7.2 B. Practice . . . 496
4.6.7.3 C. Going deeper . . . 498

4.7 Soundness 500
4.7.1 What must be proved, and why it is not obvious . . . 500
4.7.2 Derivations of formulas . . . 501
4.7.3 The context of a line . . . 502
4.7.4 The soundness theorem . . . 505
4.7.5 Every rule, checked against the tables . . . 508
4.7.6 Refuting with one row . . . 510
4.7.7 Test the subgoal . . . 512
4.7.8 Exercises . . . 512
4.7.8.1 A. Check your understanding . . . 512
4.7.8.2 B. Practice . . . 514
4.7.8.3 C. Going deeper . . . 518
4.7.9 Optional excursion: reductio says something new . . . 523

4.8 Completeness 525
4.8.1 The question soundness leaves open . . . 526
4.8.2 Bookkeeping: adding premises and chaining derivations . . . 526
4.8.3 The rules prove their own truth tables . . . 527
4.8.4 One row, one derivation . . . 530
4.8.5 Gluing the rows: splitting on an atom . . . 533
4.8.6 The completeness theorem . . . 534
4.8.7 Two columns that agree . . . 539
4.8.8 Two roads to every answer . . . 540
4.8.9 A truth-table step: TAUT . . . 541
4.8.10 Exercises . . . 543
4.8.10.1 A. Check your understanding . . . 543
4.8.10.2 B. Practice . . . 544
4.8.10.3 C. Going deeper . . . 549
4.8.10.4 R. Review . . . 555

4.9 Normal Forms and Adequate Connectives 556
4.9.1 Conjunctions and disjunctions of many formulas . . . 557
4.9.2 Literals, clauses and the two normal forms . . . 557
4.9.3 Reading a DNF off the truth table . . . 559
4.9.4 Conjunctive normal form . . . 560
4.9.5 Reading validity off a CNF . . . 563

4.9.6 Calculating with equivalences . . . 564
4.9.7 Stronger and weaker parts . . . 566
4.9.8 Adequate sets of connectives . . . 568
4.9.9 Exercises . . . 571
4.9.9.1 A. Check your understanding . . . 571
4.9.9.2 B. Practice . . . 572
4.9.9.3 C. Going deeper . . . 573

4.10 Compactness and Maximal Consistent Sets 579
4.10.1 Infinitely many premises . . . 579
4.10.2 Listing all formulas . . . 580
4.10.3 Lindenbaum's construction . . . 581
4.10.4 A row from a maximal consistent collection . . . 582
4.10.5 Completeness again . . . 584
4.10.6 Compactness . . . 584
4.10.7 Exercises . . . 586
4.10.7.1 A. Check your understanding . . . 586
4.10.7.2 B. Practice . . . 587
4.10.7.3 C. Going deeper . . . 588

5 Chapter 5 590

Predicate Logic and the Quantifier Rules 590
5.0.1 Roadmap . . . 590
5.0.2 Named moves . . . 591
5.0.3 Summary . . . 591
5.0.4 Terms . . . 591

5.1 Predicates and Quantifiers 592
5.1.1 An argument the truth table rejects . . . 592
5.1.2 Predicates: statements with holes . . . 593
5.1.3 Quantifiers and the domain . . . 594
5.1.3.1 Quantifiers in a small world . . . 595
5.1.3.2 The same sentence over different domains . . . 596
5.1.3.3 Reading a quantified sentence . . . 597
5.1.4 Bounded quantifiers . . . 598
5.1.5 The order of quantifiers . . . 599
5.1.6 Words that hide quantifiers . . . 601
5.1.7 Keys and locks . . . 602
5.1.8 What comes next . . . 603
5.1.9 Exercises . . . 603
5.1.9.1 A. Check your understanding . . . 603
5.1.9.2 B. Practice . . . 604
5.1.9.3 C. Going deeper . . . 606

5.2 Terms, Formulas, Free Variables and Substitution 608
5.2.1 Which letters are holes . . . 608
5.2.2 Terms and formulas . . . 608
5.2.3 Free and bound: drawing the arrows . . . 612
5.2.4 Substitution, by example first . . . 615
5.2.5 Capture, and the practical rule . . . 617

5.2.6 Renaming bound variables . . . 620
5.2.7 How substitution looks in a mathematical proof . . . 621
5.2.8 Exercises . . . 622
5.2.8.1 A. Check your understanding . . . 622
5.2.8.2 B. Practice . . . 624
5.2.8.3 C. Going deeper . . . 626

5.3 Using "For Every": Instantiation 628
5.3.1 Applying a statement about everything . . . 628
5.3.2 The rule . . . 629
5.3.3 Four derivations . . . 630
5.3.4 Checking a line of instantiation . . . 632
5.3.5 How this rule looks in a mathematical proof . . . 633
5.3.6 Common errors . . . 634
5.3.7 A stall . . . 635
5.3.8 The rule card . . . 635
5.3.9 The search method: one more step . . . 637
5.3.10 Exercises . . . 637
5.3.10.1 A. Check your understanding . . . 637
5.3.10.2 B. Practice . . . 639
5.3.10.3 C. Going deeper . . . 643

5.4 Proving "For Every": Let x Be Arbitrary 645
5.4.1 One number we know nothing about . . . 645
5.4.2 The flagged box . . . 646
5.4.2.1 A first flagged box: how the rule looks in a mathematical proof . . . 646
5.4.2.2 Reading the rule . . . 647
5.4.2.3 Why each clause . . . 648
5.4.3 Derivations with flagged boxes . . . 650
5.4.4 Every element of the empty set . . . 656
5.4.5 Common errors . . . 657
5.4.6 The rule card . . . 658
5.4.7 The search method: one more step . . . 659
5.4.8 Exercises . . . 660
5.4.8.1 A. Check your understanding . . . 660
5.4.8.2 B. Practice . . . 662
5.4.8.3 C. Going deeper . . . 668

5.5 Proving "There Exists": Exhibit a Witness 671
5.5.1 Showing that something exists . . . 671
5.5.2 The rule of existential introduction . . . 672
5.5.3 From "for every" to "there exists" . . . 674
5.5.4 An existential goal: postpone the witness . . . 675
5.5.5 How this rule looks in a mathematical proof . . . 678
5.5.6 The rule card . . . 679
5.5.7 Exercises . . . 680
5.5.7.1 A. Check your understanding . . . 680
5.5.7.2 B. Practice . . . 682
5.5.7.3 C. Going deeper . . . 685

5.6 Using "There Exists": Name a Witness 687
5.6.1 Using a witness whose identity is unknown . . . 688

5.6.2 The rule of existential elimination . . . 688
5.6.2.1 A first witness: $n + 2$ is even . . . 689
5.6.2.2 One box instead of many cases . . . 690
5.6.3 Using an existential: open the box early . . . 691
5.6.4 Three ways to break the rule . . . 693
5.6.5 Moving one quantifier past another . . . 699
5.6.6 Two witnesses . . . 704
5.6.7 A contradiction trapped in a box . . . 705
5.6.8 How this rule looks in a mathematical proof . . . 707
5.6.9 The four quantifier rules on one card . . . 708
5.6.10 Exercises . . . 710
5.6.10.1 A. Check your understanding . . . 710
5.6.10.2 B. Practice . . . 712
5.6.10.3 C. Going deeper . . . 717

5.7 Equality 721
5.7.1 What the rules so far cannot do with an equation . . . 721
5.7.2 The rules for equality . . . 722
5.7.3 Symmetry, transitivity and congruence . . . 724
5.7.3.1 Equal sets have the same elements . . . 727
5.7.4 Calculations are chains of substitutions . . . 727
5.7.5 What equality sounds like in a proof . . . 729
5.7.6 The search method: one more step . . . 729
5.7.7 The rule, phrase and Lean table, completed . . . 730
5.7.8 Exercises . . . 731
5.7.8.1 A. Check your understanding . . . 731
5.7.8.2 B. Practice . . . 733
5.7.8.3 C. Going deeper . . . 737

5.8 Unique Existence, Counting and Descriptions 738
5.8.1 Unique existence . . . 738
5.8.2 How uniqueness is used and proved . . . 740
5.8.3 Counting with equality . . . 743
5.8.4 The x such that φ . . . 744
5.8.5 A function symbol and its graph . . . 745
5.8.6 Definitions add words, not facts . . . 745
5.8.7 Exercises . . . 747
5.8.7.1 A. Check your understanding . . . 747
5.8.7.2 B. Practice . . . 748
5.8.7.3 C. Going deeper . . . 752
5.8.7.4 R. Review . . . 756

6 Chapter 6 759

Working with Quantifiers: Translation, Laws and Search 759
6.0.1 Roadmap . . . 760
6.0.2 Named moves . . . 760
6.0.3 Summary . . . 760
6.0.4 Terms . . . 761

6.1 Translating Mathematics: Nested and Bounded Quantifiers 761

6.1.1 Four pieces in one sentence . . . 761
6.1.2 Translation from the outside in . . . 762
6.1.3 Five traps . . . 764
6.1.4 Rules for bounded quantifiers . . . 767
6.1.5 A limit derivation with bounds . . . 771
6.1.6 Exercises . . . 774
6.1.6.1 A. Check your understanding . . . 774
6.1.6.2 B. Practice . . . 775
6.1.6.3 C. Going deeper . . . 778

6.2 Negating and Moving Quantifiers 781
6.2.1 Not every, and some not . . . 782
6.2.2 The negation of an implication . . . 784
6.2.3 Bounded quantifiers keep their bounds . . . 785
6.2.4 Negating a definition, mechanically . . . 787
6.2.5 Four forms and their square . . . 792
6.2.6 How the laws look in a mathematical proof . . . 793
6.2.7 Moving a quantifier across a connective . . . 794
6.2.8 Prenex form . . . 800
6.2.9 Exercises . . . 800
6.2.9.1 A. Check your understanding . . . 800
6.2.9.2 B. Practice . . . 802
6.2.9.3 C. Going deeper . . . 807

6.3 Finding First-Order Derivations 811
6.3.1 The search method, extended . . . 811
6.3.2 Three moves for quantifiers . . . 812
6.3.3 Steps that are right for propositional reasons . . . 815
6.3.4 The propositional core . . . 818
6.3.5 Four kinds of task, with quantifiers . . . 819
6.3.6 A library of first-order sequents . . . 822
6.3.7 Exercises . . . 823
6.3.7.1 A. Check your understanding . . . 823
6.3.7.2 B. Practice . . . 823
6.3.7.3 C. Going deeper . . . 829

6.4 Uniformly Continuous Implies Continuous, Rule by Rule 831
6.4.1 Two definitions that differ in one place . . . 831
6.4.2 The skeleton first . . . 832
6.4.3 The derivation . . . 832
6.4.4 Reading dependence off the boxes . . . 834
6.4.5 The converse breaks at a side condition . . . 835
6.4.6 A function that is continuous and not uniformly continuous . . . 836
6.4.7 Witness functions . . . 838
6.4.8 Exercises . . . 840
6.4.8.1 A. Check your understanding . . . 840
6.4.8.2 B. Practice . . . 842
6.4.8.3 C. Going deeper . . . 845
6.4.8.4 R. Review . . . 848

7 Chapter 7 **851**

Models and Theories **851**
7.0.1 Roadmap . . . 852
7.0.2 Named moves . . . 852
7.0.3 Summary . . . 853
7.0.4 Terms . . . 853

7.1 Structures and Satisfaction **853**
7.1.1 Why a row is not enough . . . 854
7.1.2 Structures . . . 854
7.1.3 Drawing a finite structure . . . 856
7.1.4 Assignments and the value of a term . . . 857
7.1.5 Satisfaction . . . 858
7.1.6 Only the free variables matter . . . 860
7.1.7 Semantic entailment . . . 862
7.1.8 Proving an entailment by meaning . . . 863
7.1.9 Why domains are non-empty . . . 864
7.1.10 Exercises . . . 865
7.1.10.1 A. Check your understanding . . . 865
7.1.10.2 B. Practice . . . 867
7.1.10.3 C. Going deeper . . . 869

7.2 Counter-models **870**
7.2.1 Three pictures . . . 870
7.2.2 Counter-models . . . 871
7.2.3 Finding a counter-model . . . 873
7.2.4 Small domains, by expansion . . . 875
7.2.5 Each side condition, broken . . . 876
7.2.6 From counter-models to non-derivability . . . 877
7.2.7 Three classes of sentences . . . 878
7.2.8 Extensionality: an axiom, not a law of logic . . . 880
7.2.9 Exercises . . . 881
7.2.9.1 A. Check your understanding . . . 881
7.2.9.2 B. Practice . . . 882
7.2.9.3 C. Going deeper . . . 887

7.3 Soundness for Predicate Logic **891**
7.3.1 What carries over, and what is new . . . 891
7.3.2 Contexts, once more . . . 893
7.3.3 Substitution and satisfaction . . . 894
7.3.4 The soundness theorem . . . 896
7.3.5 Counter-models prove non-derivability . . . 899
7.3.6 Every side condition is a case of the proof . . . 900
7.3.7 Exercises . . . 901
7.3.7.1 A. Check your understanding . . . 901
7.3.7.2 B. Practice . . . 902
7.3.7.3 C. Going deeper . . . 903

7.4 Theories and Their Models **905**
7.4.1 Theories and their models . . . 905
7.4.2 A gallery of theories . . . 906

7.4.3 A theorem of a theory . . . 910
7.4.4 Consistency . . . 911
7.4.5 Independence . . . 912
7.4.6 Exercises . . . 913
7.4.6.1 A. Check your understanding . . . 913
7.4.6.2 B. Practice . . . 915
7.4.6.3 C. Going deeper . . . 917

7.5 Definability and Isomorphism 920
7.5.1 Formulas define sets . . . 921
7.5.2 ∃ is a projection . . . 922
7.5.3 Define it by a formula . . . 923
7.5.4 Isomorphic structures (second pass) . . . 924
7.5.5 What an automorphism cannot move (second pass) . . . 927
7.5.6 Exercises . . . 928
7.5.6.1 A. Check your understanding . . . 928
7.5.6.2 B. Practice . . . 930
7.5.6.3 C. Going deeper . . . 931
7.5.6.4 R. Review . . . 932

7.6 What First-Order Logic Can and Cannot Say 936
7.6.1 Completeness . . . 936
7.6.2 Compactness . . . 938
7.6.3 Large models and small ones . . . 939
7.6.4 Numbers that are not numerals . . . 939
7.6.5 Induction: a schema and an axiom . . . 940
7.6.6 What can be decided . . . 941
7.6.7 Exercises . . . 942
7.6.7.1 A. Check your understanding . . . 942
7.6.7.2 B. Practice . . . 943
7.6.7.3 C. Going deeper . . . 945

8 Chapter 8 947

From Derivations to Mathematical Prose 947
8.0.1 Roadmap . . . 949
8.0.2 Named moves . . . 949
8.0.3 Summary . . . 950
8.0.4 Terms . . . 950

8.1 The Connective Rules in Prose 950
8.1.1 The rule card . . . 951
8.1.2 Which steps may go unsaid . . . 952
8.1.3 The same proof, twice . . . 953
8.1.4 How to read the dictionary . . . 954
8.1.5 Conjunction . . . 954
8.1.6 Implication . . . 955
8.1.7 Disjunction and cases . . . 956
8.1.8 Negation and contradiction . . . 957
8.1.9 Proof by contradiction . . . 959
8.1.10 The biconditional . . . 960

8.1.11 Equivalent conditions, by a cycle . . . 961
8.1.12 The three faces of "by contradiction" . . . 962
8.1.13 Costumes . . . 963
8.1.14 Exercises . . . 965
8.1.14.1 A. Check your understanding . . . 965
8.1.14.2 B. Practice . . . 967
8.1.14.3 C. Going deeper . . . 969

8.2 The Quantifier Rules in Prose 971
8.2.1 Three sentences, three rules . . . 971
8.2.2 The quantifier rows . . . 973
8.2.3 The dependency parenthesis . . . 974
8.2.4 Equality and unique existence . . . 975
8.2.5 Definitions, descriptions and sets . . . 977
8.2.6 Induction, lemmas and symmetry . . . 978
8.2.6.1 It suffices to show . . . 979
8.2.6.2 Induction . . . 980
8.2.6.3 Without loss of generality . . . 980
8.2.7 What the sentence does not say . . . 981
8.2.8 Exercises . . . 982
8.2.8.1 A. Check your understanding . . . 982
8.2.8.2 B. Practice . . . 984
8.2.8.3 C. Going deeper . . . 986

8.3 Four Levels of One Proof 987
8.3.1 One short proof, four texts . . . 988
8.3.2 Which steps may be silent . . . 990
8.3.3 The irrationality of the square root of two, four times . . . 993
8.3.4 Uniform continuity, twice . . . 996
8.3.5 It suffices to show, as a column . . . 999
8.3.6 Expand to check . . . 1000
8.3.7 Exercises . . . 1002
8.3.7.1 A. Check your understanding . . . 1002
8.3.7.2 B. Practice . . . 1004
8.3.7.3 C. Going deeper . . . 1006

8.4 Reading a Proof: Self-Explanation 1010
8.4.1 Two readers of one proof . . . 1010
8.4.2 The training passage . . . 1011
8.4.3 How to explain a line . . . 1012
8.4.4 Naming the unstated fact . . . 1013
8.4.5 Strategies of good readers . . . 1014
8.4.6 Exercises . . . 1015
8.4.6.1 A. Check your understanding . . . 1015
8.4.6.2 B. Practice . . . 1016
8.4.6.3 C. Going deeper . . . 1019

8.5 Writing a Proof 1021
8.5.1 A procedure . . . 1021
8.5.2 The procedure, run once . . . 1022
8.5.2.1 Scratch paper and the page . . . 1026
8.5.3 Conventions of written mathematics . . . 1027

8.5.4 Matching sentences and rules . . . 1028
8.5.5 Generalizing a proof . . . 1030
8.5.6 Exercises . . . 1031
8.5.6.1 A. Check your understanding . . . 1031
8.5.6.2 B. Practice . . . 1033
8.5.6.3 C. Going deeper . . . 1037

8.6 Checking Proofs, and What Is Not Allowed 1040
8.6.1 Two questions for every line . . . 1041
8.6.2 The catalogue . . . 1042
8.6.3 Four proofs, checked . . . 1046
8.6.4 Exercises . . . 1049
8.6.4.1 A. Check your understanding . . . 1049
8.6.4.2 B. Practice . . . 1051
8.6.4.3 C. Going deeper . . . 1053
8.6.4.4 R. Review . . . 1057

8.7 The Dispute over Excluded Middle 1058
8.7.1 Irrational powers, with and without a witness . . . 1058
8.7.2 What a proof provides . . . 1060
8.7.3 One extra rule . . . 1061
8.7.4 Excluded middle cannot be refuted . . . 1062
8.7.5 Which laws survive without RAA . . . 1064
8.7.6 A short history . . . 1065
8.7.7 What analysis would lose . . . 1066
8.7.8 Proofs as programs . . . 1067
8.7.9 The book's position . . . 1067
8.7.10 Exercises . . . 1068
8.7.10.1 A. Check your understanding . . . 1068
8.7.10.2 B. Practice . . . 1069
8.7.10.3 C. Going deeper . . . 1074
8.7.10.4 R. Review . . . 1076

8.8 Stages of Knowledge: Kripke Models 1080
8.8.1 Stages of knowledge . . . 1080
8.8.2 Forcing . . . 1081
8.8.3 A first countermodel . . . 1081
8.8.4 Knowledge persists . . . 1082
8.8.5 The rules without RAA are sound for stages . . . 1083
8.8.6 Two branches . . . 1085
8.8.7 What two or three stages settle . . . 1086
8.8.8 Searching for a countermodel . . . 1086
8.8.9 Exercises . . . 1087
8.8.9.1 A. Check your understanding . . . 1087
8.8.9.2 B. Practice . . . 1088
8.8.9.3 C. Going deeper . . . 1092

Preface

This volume is for a reader who knows school algebra and has never been asked to write a proof. It teaches one skill: to read a mathematical statement exactly, and to prove it in sentences that a reader can check line by line.

It teaches logic as the grammar of proof. A proof is first written as a **derivation**, a chain in which every line is a premise, an assumption, or one application of a rule, so that checking it needs no idea: for each line we ask which rule produced it and from which lines. Each rule is introduced beside the sentence a mathematician writes when using it, and it is practiced at once, in short derivations and in proofs about numbers. The last chapter turns derivations back into ordinary mathematical prose, and shows that the rules are still there underneath.

What it covers.

Chapter	What it does
1	The format of a derivation, and the rules for "and", "implies" and "or"
2	The rules for "not" and "if and only if", proof by contradiction and the classical rules, and a method for finding derivations
3	Truth tables, induction on $\mathbb{N}$, and the theorems that the rules are sound and complete
4	"For every" and "there exists": terms and formulas, the quantifier rules and their side conditions, equality, and models
5	From derivations to mathematical prose: four levels of writing, habits for reading, writing and checking proofs, and a catalogue of the errors that recur

How to read it. Read the chapters in order: each uses the ones before it. Every section has Quick checks with folded answers; answer each one before you unfold it. The exercises of a section come in three groups, A to check understanding, B for practice and C to go deeper, and the last section of each chapter adds R items that reach back to earlier chapters. Do the derivations on paper, line by line, before reading the solution. A chapter's opening page lists what it needs, with links, and the time it takes. Second-reading folds hold edge cases and longer justifications; they can be left closed on a first reading, but they are not optional, since later proofs may use what they contain. Sections and results marked optional can be skipped, and nothing outside them depends on them.

Why rules before truth tables. Many logic books begin with truth tables, where the row

"a false hypothesis makes $P \to Q$ true" has to be accepted as a convention. This book starts from the rules. Section 1.1 reads that row informally, as a promise that is not broken; Chapter 2 derives it from the rules, and Chapter 4 brings in the tables and proves that rules and tables agree.

What comes next. *Sets, Relations, and Cardinality* puts the logic to work on sets, functions, relations, induction and the sizes of infinite sets. Its third chapter ends with Gate 0, an examination on both volumes, and a reader who passes it is ready for *Analysis on the Real Line*. *Number Systems: From* $\mathbb{N}$ *to* $\mathbb{R}$ builds the numbers themselves out of sets.

Chapter 1

Connectives and Derivations

Section 1.1 opens with two arguments that every integer with an even square is even. Every claim in both is true, and yet only one is a proof. The other, Argument B, starts from "n is even", reaches "n^2 is even", and then announces the reverse. Spotting the flaw once it is pointed out is easy. Naming it, so that the same flaw can be caught in an argument we have never seen, needs an answer to a harder question: when is a list of true claims a proof? The motivating problem of this chapter is to *name exactly what is wrong with Argument B*, and the answer arrives in Section 1.4.

This chapter builds the calculus that answers it: a proof is a chain in which every line is a premise, an assumption, or one application of a rule. It gives the rules for "and", "implies" and "or", one at a time; Chapter 2 adds "not", "iff" and proof by contradiction.

Before you start: five tempting misreadings. Each is a misreading that this chapter corrects; decide each now, and check again when you reach the section named.

Quick check.

1. Let x be a real number. "If $x > 3$, then $x^2 > 9$" is false at $x = 1$, since $x > 3$ fails there.
2. Let n be an integer. "n is divisible by 4 only if n is even" says that every even number is divisible by 4.
3. In a proof that begins "Suppose $x > 2$" and derives $x^2 > 4$, the line $x^2 > 4$ may be used anywhere later.
4. "$5 \geq 0$" is false, because $5 \neq 0$.
5. Let x be a real number. From "$x \geq 1$ or $x \leq 2$" and "$x \geq 1$" it follows that $x \leq 2$ fails.

Solution.

1. No. An implication breaks only when its hypothesis holds and its conclusion fails; at $x = 1$ it promises nothing (Section 1.1, "When is an implication broken?").
2. No. "P only if Q" is $P \to Q$: divisible by 4 implies even, which is true. The claim reads the converse, false at $n = 2$ (Section 1.1).
3. No. It holds only inside the supposition's box. Outside it, only "if $x > 2$, then $x^2 > 4$" may be used, and $x^2 > 4$ fails at $x = 0$ (Section 1.4).
4. No. $5 \geq 0$ is $(5 > 0) \vee (5 = 0)$, and an inclusive "or" needs one true side: $5 > 0$ is true (Section 1.5).

5. No. At $x = \frac{3}{2}$ both sides hold. The form "from $P \vee Q$ and P, conclude not Q" is affirming a disjunct, a bad form (Sections 1.1 and 1.6).

What you need. School algebra only: the rules of signs, adding and multiplying inequalities, and the facts that every integer is even or odd and every real number is non-negative or negative. They are used as facts in examples; *Analysis on the Real Line*, Chapter 1 proves the order facts from the axioms for the real numbers. No earlier chapter is needed.

Time. About 40 hours on the core track (all of A, one B and one C per section, the exit checks); about 65 hours for everything.

Minimum before *Sets, Relations, and Cardinality.* Chapters 1–6 and 8, core; Sections 7.1–7.2 before its Chapter 3.

1.0.1 Roadmap

- **Statements, Arguments and How to Break One.** Statements and the connectives, a phrase book for the English that hides them, the promise reading of "if ... then", and how to break a bad argument form by making its counterexample set true.
- **Derivations, Starting with And.** The format of a derivation, with numbered lines, premises and a justification on every line, and the first two rules of logic, $\wedge$I and $\wedge$E.
- **Using an Implication: Modus Ponens.** The rule $\to$E, and chains of implications found forwards and backwards.
- **Proving an Implication: the Rule of the Box.** The rule $\to$I, the box and the scope rule for citing lines, and a diagnosis of exactly what Argument B did wrong.
- **Or, Introduced.** The rule $\vee$I, and the habit of postponing a disjunctive goal until one side is in hand.
- **Using Or: Proof by Cases.** The rule $\vee$E, the laws of "or" and the constructive dilemma, case proofs about absolute values, and the whole chapter's rule card.

1.0.2 Named moves

- **Try to break it.** To test a form, put statements in for the letters that make every premise true and the conclusion false; one such choice shows the form is bad. (Section 1.1)
- **Two ways to attack an argument.** Show a premise false, or show a step's form bad; the two attacks are independent. (Section 1.1)
- **Read the skeleton off the goal.** The goal's main connective chooses the introduction rule: $\wedge$ gives two subgoals, $\to$ a box, $\vee$ a choice to postpone. (Section 1.2)
- **Use a premise by its main connective.** Its elimination rule says what you may do with it: take a conjunction apart, feed an implication its hypothesis, split a disjunction into cases. (Section 1.2)
- **Open a box, close a box.** An assumption lives exactly as long as its box. (Section 1.4)
- **Work from both ends.** Premises at the top, the goal at the bottom; fill the gap from both sides. (Section 1.2)
- **Postpone a disjunction.** Choose a side of a goal $P \vee Q$ only when that side is in hand. (Section 1.5)
- **Cases end in the current goal.** Both boxes of $\vee$E close on the statement you are trying to reach. (Section 1.6)

The chapter's eight moves are collected again at the end of Section 1.6, beside the rules that make them.

1.0.3 Summary

A proof is a chain in which every line is a premise, an assumption, or one step of a rule, and checking one needs no idea. The chapter gave the six rules for "and", "implies" and "or", with repetition, and with them derived that "and" commutes and regroups (Proposition 1.2.1, Proposition 1.2.2), that implications chain (Proposition 1.4.1), that two hypotheses may be taken at once or one at a time (Proposition 1.4.2), that "or" commutes and regroups (Example (First derivations with "or"), Example (Regrouping a disjunction)), that "and" distributes over "or" (Proposition 1.6.1) and the constructive dilemma (Example (Constructive dilemma)). It named Argument B's error: a box opened at the wrong end proves the converse. A form is bad when some choice of statements makes its counterexample set true, and breaking a form is how a bad step is shown bad. Still open: the chapter has no rules for "not" or for contradiction, and the fact Argument B set out to prove waits for them. That is Chapter 2's question.

1.0.4 Terms

- **argument form, good and bad**: bad if some choice of statements breaks it (Definition 1.1.3).
- **assumption, discharged**: an assumption whose box has closed (Definition 1.4.1).
- **available**: citable at a line, by the scope rule (Definition 1.4.2).
- **box**: a block of lines opened by an assumption (Definition 1.4.2).
- **box-free derivation**: premises, then one rule step per line (Definition 1.2.2).
- **conjunction rules** $\wedge$I, $\wedge$E: build from both sides; take apart (Definition 1.2.5).
- **counterexample set**: the premises together with the negated conclusion (Definition 1.1.3).
- **derivable**, $\Gamma \vdash \psi$: some derivation from premises in Γ ends with ψ (Definition 1.2.3).
- **derivation**: each line a premise, assumption, repetition or rule step (Definition 1.4.2).
- **disjunction rules** $\vee$I, $\vee$E, **proof by cases**: weaken to a disjunction; split one into cases (Definition 1.6.1).
- **implication rules** $\to$I, $\to$E: prove by a box; use by modus ponens (Definition 1.4.1).
- **main connective**: the connective applied last in building a statement (Definition 1.1.2).
- **modus ponens**: from $\varphi \to \psi$ and φ, conclude ψ (Definition 1.3.1).
- **or-introduction**: from one side, conclude the disjunction (Definition 1.5.1).
- **repetition** R: copy the statement of an available line (Definition 1.4.2).
- **rule schema**: given lines above, one new line below (Definition 1.2.1).
- **sequent**, turnstile: $P_1, \ldots, P_n \vdash Q$, the claim that Q is derivable (Definition 1.2.4).
- **statement**: a declarative sentence, either true or false (Definition 1.1.1).

1.1 Statements, Arguments and How to Break One

When is an argument a proof? This section says what the sentences of a proof are, how they are built from smaller ones, and how to test the **form** of an argument by trying to break it.

1.1.1 Two arguments about squares

Recall that an integer n is **even** when $n = 2k$ for some integer k, and **odd** when $n = 2k + 1$ for some integer k. Here is a fact about integers.

> For every integer n, if n^2 is even, then n is even.

Read the following two arguments for it. Each is written in the style of a textbook proof, and each claim in each of them is true.

Argument A. Suppose n is odd, say $n = 2k + 1$ for an integer k. Then $n^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$, so n^2 is odd, and therefore n^2 is not even. So an odd n never has an even square. Since every integer is even or odd, an integer with an even square must be even.

Argument B. Suppose n is even, say $n = 2k$ for an integer k. Then $n^2 = 4k^2 = 2(2k^2)$, and $2k^2$ is an integer, so n^2 is even. Hence if n^2 is even, then n is even.

Argument A is a correct proof. Argument B is not, although none of its claims is false. It starts from "n is even" and arrives at "n^2 is even". That is a perfectly good proof of a different statement: "if n is even, then n^2 is even". The claim asks for the journey in the opposite direction, from the square back to the number, and the last sentence of B announces a conclusion that the argument never reached.

Seeing this is easy once it is pointed out. Saying in general what went wrong is harder, and that is what this chapter and the next are for. To name the error we need to know what the rules of proof are, and before that, what they act on. So we start with the smallest pieces: the sentences a proof is made of.

1.1.2 Statements

A proof is written in sentences, but not every sentence can appear in one. "Is 91 prime?" asserts nothing, so it cannot serve as a premise or a conclusion; neither can "Let us now turn to the case $n = 0$." The sentences a proof asserts are the ones that make a claim. A proof also has steps of another kind, such as "Suppose n is odd", which assert nothing but set up a temporary hypothesis; their rules come in Section 1.4.

A statement is a sentence that is either true or false.

Definition 1.1.1. Statement

A **statement** is a declarative sentence that is **either true or false**, and not both.

In words: a statement makes a claim, and the claim is either correct or incorrect. The definition asks that one of the two is the case. It does **not** ask that we know which one. A statement whose truth nobody has settled is still a statement.

Example. Statements and nonstatements

Decide which of the following are statements.

(a) 7 is a prime number.
(b) $0.999 > 1$.
(c) There are infinitely many primes p for which $p + 2$ is also prime.
(d) Is $2^{11} - 1$ prime?
(e) $x > 2$.
(f) $3 > 2$.

Solution.

(a) This is a statement, and it is true.

(b) This is a statement, and it is false. A false sentence is as much a statement as a true one.

(c) This is a statement. It is the twin prime conjecture, and at the time of writing nobody knows whether it is true. It is certainly true or false, which is all the definition asks.

(d) This is a question, so it is not a statement. The sentence "$2^{11} - 1$ is prime" is a statement, and a false one, since $2^{11} - 1 = 2047 = 23 \cdot 89$.

(e) This is not a statement. Whether it is true depends on what x is: it is true when $x = 3$ and false when $x = 0$, and as written nothing says which. Such a sentence is called an **open sentence**.

(f) This is the sentence of (e) with x replaced by 3, and it is a true statement.

Items (e) and (f) differ by one symbol, and the difference is the whole point. The open sentence $x > 2$ becomes a statement in two ways. We can put a number in for x, as in (f). Or we can say which x we mean in some other way, as in "for every real number x, $x > 2$" (a false statement) or "for some real number x, $x > 2$" (a true one). The words "for every" and "for some" are the **quantifiers**.

Remark. **Convention: a fixed but arbitrary number.** When an example says "let x be a real number" and then talks about x, the letter names **one** number, fixed for the whole example, that we are not told. Every sentence about x is then a statement, true or false according to which number x is: with $x = 3$, "$x > 2$" is true; with $x = 0$, it is false; and the example does not know which x it has. So an argument about x may use only what holds whatever x is, and an argument that never peeks at x works for $x = 3$, for $x = 0$, and for every other number at once.

Quick check. Let x be a real number. Is "$x > 2$ or $x \leq 2$" a statement? Is it true?

Solution. Yes: by the convention it is a claim about one fixed number x. And it is true whatever x is, since every real number is either greater than 2 or at most 2.

↻ **Second reading** · *Statements or propositions?*

Remark. The book reserves the words "proposition" and "theorem" for proved results, so it calls the sentences of logic **statements**. Other books call them propositions, which is where the name **propositional logic**, for what this chapter and the next study, comes from.

1.1.3 The connectives

Most statements in mathematics are built from smaller ones. "$0 < x$ and $x < 1$" is built from "$0 < x$" and "$x < 1$" with the word "and". "If n^2 is even, then n is even" is built from two statements with "if ... then". The words that do the building are the **connectives**. A statement built from smaller statements by a connective is **compound**; one that is not is **atomic**. For this chapter, "$x > 0$" is atomic: it has an inner structure, a number and an order, but no connective joins parts of it.

We write P, Q, R, S for statements whose inner structure we are ignoring. They stand for **arbitrary** statements, atomic or compound. There are five connectives. In the last column, x is a real number, and every example is true whatever x is.

Symbol	Name	Read	Example
$P \wedge Q$	conjunction	P and Q	$x^2 \geq 0$ and $x^2 + 1 > 0$
$P \vee Q$	disjunction	P or Q	$x \geq 0$ or $x < 0$
$P \to Q$	implication	if P, then Q	if $x > 1$, then $x^2 > 1$
$P \leftrightarrow Q$	biconditional	P if and only if Q	$x^2 = 0$ if and only if $x = 0$
$\neg P$	negation	not P	not $x^2 < 0$, that is, $x^2 \geq 0$

To the five connectives we add one statement with no parts at all, written $\perp$ and read "contradiction". It stands for a statement that is false however things stand: $0 = 1$ will do. A proof that reaches $\perp$ has shown that the situation it was working in cannot occur. That is the heart of proof by contradiction, and the rules for $\perp$ come with those for "not".

In $P \to Q$, the statement P is the **hypothesis** (or antecedent) and Q the **conclusion** (or consequent). "If and only if" is often shortened to **iff**.

Several English words become $\wedge$: "but", "although", and "while" all assert both parts, and add a tone of contrast that logic ignores. "2 is prime but not odd" is a conjunction. So is "neither ... nor", in disguise: "9 is neither even nor prime" asserts that 9 is not even **and** that 9 is not prime, which is $\neg P \wedge \neg Q$ with P: "9 is even" and Q: "9 is prime".

Chained inequalities are conjunctions in disguise too. $0 < x < 1$ abbreviates $(0 < x) \wedge (x < 1)$, and $a \leq b \leq c$ abbreviates $(a \leq b) \wedge (b \leq c)$. The symbol $\leq$ itself hides a disjunction: $a \leq b$ means $(a < b) \vee (a = b)$. And $a \neq b$ is the negation $\neg(a = b)$. All three facts are used often in the sections that follow.

In theorem statements the book writes $\Rightarrow$ and $\iff$, as mathematicians usually do, and in running prose it says "implies" and "if and only if". The arrows $\to$ and $\leftrightarrow$ are kept for statements written in logical symbols, as in this section, whether built from letters like $P \to Q$ or from formulas like $(x > 1) \to (x^2 > 1)$, and for the formal derivations that start in the next. Nothing but typography separates the two pairs.

Warning. In everyday English, "or" is often exclusive: "soup or salad" on a menu means one but not both. In mathematics, $P \vee Q$ is **inclusive**: it is true when P is true, when Q is true, and when both are. For a real number x, "$x \geq 0$ or $x \leq 0$" is true whatever x is; at $x = 0$ both parts are true, and that does not spoil it. An exclusive "or" must be said in full: "P or Q, but not both", which is $(P \vee Q) \wedge \neg(P \wedge Q)$.

Warning. The word "if" means different things in a theorem and in a definition. In a theorem, "if P, then Q" is one direction only: "if n is divisible by 4, then n is even" is true, and its converse is false at $n = 2$. In a definition, "a number is **even** if it is twice an integer" is meant as "if and only if": a number that is not twice an integer is not called even.

1.1.4 When is an implication broken?

Read $P \to Q$ as a **promise**: whenever P holds, Q holds too. A promise is broken only when its condition is met and the promised thing fails.

An implication is broken only when its hypothesis holds and its conclusion fails.

Take the table's example, "if $x > 1$, then $x^2 > 1$". Its two parts can be true or false in four combinations:

Table 1.1.2: Informal: a reading of "if", not yet a rule of proof.

P: $x > 1$	Q: $x^2 > 1$	the promise	a value of x
true	true	kept	$x = 3$
true	false	**broken**	none: this is why the promise holds for every x
false	true	not broken: it says nothing	$x = -3$
false	false	not broken: it says nothing	$x = 0$

(No value lands in the second row: if $x > 1$, then $x^2 > x > 1$.) This answers the question "what if P is false?": then $P \to Q$ is not broken, so it counts as true. To show $P \to Q$ false, exhibit one case with P true and Q false.

Quick check. Which of $x = 3$, $x = -3$, $x = 1$ breaks "if $x^2 > 4$, then $x > 2$"?

Solution. Only $x = -3$: $9 > 4$ holds and $-3 > 2$ fails. At $x = 3$ both parts hold; at $x = 1$ the hypothesis fails, so nothing is promised.

1.1.5 "Only if", "necessary" and "sufficient"

English has many ways of saying $P \to Q$: "P implies Q", "Q if P", "Q whenever P", "P only if Q", "P is sufficient for Q", and "Q is necessary for P". To sort out the last three, ask of each sentence: *when is it broken?* Let x be a real number, with P: "$x > 2$" and Q: "$x^2 > 4$".

- "$x > 2$ only if $x^2 > 4$" says that $x > 2$ never happens without $x^2 > 4$. It is broken exactly when $x > 2$ and $x^2 \leq 4$, which is exactly when $P \to Q$ is broken. So "P only if Q" is $P \to Q$. It is true here: if $x > 2$, then $x^2 > 2x > 4$.
- The reverse, "$x^2 > 4$ only if $x > 2$", is $Q \to P$, broken by $x = -3$ (the Quick check above).

- "$x^2 > 4$ is **necessary** for $x > 2$" says that there is no $x > 2$ without $x^2 > 4$: again $P \to Q$. "$x > 2$ is **sufficient** for $x^2 > 4$" says that $x > 2$ is enough to guarantee $x^2 > 4$: $P \to Q$ once more.

English	broken when	symbols
P only if Q	P true, Q false	$P \to Q$
Q if P	P true, Q false	$P \to Q$
Q is necessary for P	P true, Q false	$P \to Q$
P is sufficient for Q	P true, Q false	$P \to Q$

"If" and "only if" put the same two statements in opposite directions: "Q if P" is $P \to Q$, while "Q only if P" is $Q \to P$. Confusing them is Argument B's mistake.

1.1.6 Translating a sentence

Translating into symbols is a skill of its own, best learned slowly, in three steps. Find the atomic statements, the smallest pieces that make a claim, and give each a letter. Find the connectives, whatever English words carry them. Then decide how the pieces group.

Take "if 12 is divisible by 3 and by 4, then 12 is divisible by 6". The atomic statements are P: "12 is divisible by 3", Q: "12 is divisible by 4" and R: "12 is divisible by 6". English lets "12 is divisible by" be said once and shared; logic writes each piece in full. The words "if ... then" carry $\to$, and "and" carries $\wedge$. The hypothesis is everything between "if" and "then", so the statement is $(P \wedge Q) \to R$, not $P \wedge (Q \to R)$. The second reading would claim P outright, and the English sentence claims nothing outright: it only says what follows from what.

A second example shows a connective hidden inside a word. Let x be a real number, and read "x is non-zero, and if x is positive, then so is x^3". With Z: "$x = 0$", P: "$x > 0$" and C: "$x^3 > 0$", "non-zero" is $\neg Z$, and the whole sentence is $\neg Z \wedge (P \to C)$. Letters that recall their statements, as here, often help; a letter is only a name.

In both examples the English word order is a poor guide to the grouping. When in doubt, ask of each connective: which two statements does it join?

1.1.7 A phrase book

Many English phrases carry a connective in disguise. The table collects the common ones; in each row, ask the question of "Only if": when is the sentence broken? In the examples, n is an integer and x is a real number.

English	paraphrase	symbols	example
P provided Q	if Q, then P	$Q \to P$	$x^2 > 4$ provided $x > 2$
P whenever Q	if Q, then P	$Q \to P$	n^2 is even whenever n is even
P but Q	P and Q	$P \wedge Q$	2 is prime but even
neither P nor Q	not P, and not Q	$\neg P \wedge \neg Q$	9 is neither even nor prime
not both P and Q	not (P and Q)	$\neg(P \wedge Q)$	9 is not both odd and prime

English	paraphrase	symbols	example
exactly one of P, Q	P or Q, and not both	$(P \vee Q) \wedge \neg(P \wedge Q)$	exactly one of $x > 0$, $x < 1$
at least two of P, Q, R	some two of them hold	$((P \wedge Q) \vee (P \wedge R)) \vee (Q \wedge R)$	at least two of: n even, $n > 0$, n a multiple of 3

"Unless" is Exercise C1.

"Neither ... nor" and "not both" differ by one word and say different things. At $n = 3$, "n is neither even nor prime" is false, since 3 is prime, while "n is not both even and prime" is true. In the "exactly one" row, at $x = 2$ exactly one part holds ($2 > 0$, and $2 < 1$ fails); at $x = \frac{1}{2}$ both hold, so "exactly one" fails. In the last row, $n = 6$ satisfies all three parts, and $n = -3$ only one: it is a multiple of 3, but odd and negative.

Quick check. Let n be an integer, with E: "n is even" and P: "n is prime". Translate "exactly one of: n is even, n is prime", and decide it at $n = 2$, $n = 9$ and $n = 7$.

Solution. $(E \vee P) \wedge \neg(E \wedge P)$. It is false at 2, where both parts hold; false at 9, where neither does; and true at 7, which is prime and not even.

Not every "and" joins two statements. Some sentences with "and" in them make one claim about several things together.

- "4 and 9 are coprime." *Not:* "4 is coprime $\wedge$ 9 is coprime". Coprime is a relation between two numbers (their only common positive divisor is 1), so the sentence is one atomic statement about a pair. Contrast "4 and 9 are squares", which is a conjunction.
- "The points A, B, C are collinear." *Not:* a conjunction of three statements about single points. Collinearity is one statement about all three points at once.
- "1.414 is an approximate root of $x^2 - 2 = 0$." *Not:* "1.414 is approximate $\wedge$ 1.414 is a root". It is not a root, since $1.414^2 = 1.999396$. "Approximate" changes the meaning of "root", and nothing here is a conjunction.

By contrast, "12 is a positive even number" is the conjunction of "$12 > 0$" and "12 is even".

Remark. **Everyday readings.** Four kinds of word mean something different in mathematics from what they mean in conversation, and the everyday reading translates wrongly.

- "Or" is inclusive: the first warning of "The connectives" above.
- In a definition, "if" is read more strongly, as "if and only if": the second warning there.
- "Some" never says "not all". "Some solutions of $x^2 = x$ are non-negative" is true, although every solution (0 and 1) is non-negative.
- A word without a sharp boundary makes no statement. "x is large" is not a statement until a bound is fixed, as in "$x > 10^6$".

1.1.8 Reading a statement: precedence and the main connective

Connectives can be combined without limit, and a string of symbols is only useful if it can be read in one way. Consider $\neg P \wedge Q$. Does it deny P and then assert Q, or does it deny the whole of $P \wedge Q$? These are different claims. With P: "n is even" and Q: "$n > 0$", the first says that n is odd and positive, while the second says only that n is not a positive even number, which $n = -4$ satisfies.

Parentheses settle every such question, and a few conventions let us drop the ones that would only add clutter.

(1) $\neg$ binds most tightly: it applies to the smallest complete statement that follows it. So $\neg P \wedge Q$ means $(\neg P) \wedge Q$, and to deny $P \wedge Q$ we write $\neg(P \wedge Q)$.

(2) $\wedge$ and $\vee$ bind more tightly than $\to$ and $\leftrightarrow$. So $P \wedge Q \to R$ means $(P \wedge Q) \to R$.

In a mixture of $\wedge$ and $\vee$, or of $\to$ and $\leftrightarrow$, the parentheses are always written.

Remark. When parentheses in a chain of arrows are dropped, they group to the right: $P \to Q \to R$ means $P \to (Q \to R)$. This chapter always writes them.

In this chapter we parse by eye.

The conventions only decide where parentheses may be dropped. The next definition names the one piece of information that reading a statement is for.

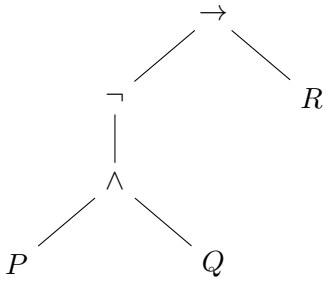
Definition 1.1.2. Main Connective

The **main connective** of a compound statement is the connective applied **last** when the statement is built up from its parts. An atomic statement has no main connective.

The main connective is found by undoing the construction from the outside. Take

$$\neg(P \wedge Q) \to R.$$

By convention (1), the $\neg$ applies only to the parenthesized $P \wedge Q$. That leaves a single connective outside the parentheses, the $\to$, joining $\neg(P \wedge Q)$ on its left and R on its right. So the last step joined $\neg(P \wedge Q)$ and R with $\to$, and $\to$ is the main connective. Its hypothesis $\neg(P \wedge Q)$ has main connective $\neg$, applied to $P \wedge Q$, whose main connective is $\wedge$. The construction can be drawn as a tree, with the main connective at the root and the letters at the leaves.



Reading the tree from the top: the statement is an implication; its hypothesis is a negation; what is negated is a conjunction. Reading from the bottom retraces how the statement was

built. A one-symbol change moves the root: in $\neg(P \wedge Q \to R)$ the parentheses make the whole implication the thing denied, and the main connective is $\neg$.

The main connective is not grammar for its own sake. It is the first thing to find when a proof has to be written, because it decides how the proof begins. To prove "if $0 < x < 1$, then $x^2 < x$", whose main connective is $\to$, a proof starts "Suppose $0 < x < 1$." To prove "$x^2 + 1 > 0$ and $x^2 + 1 \neq 0$", whose main connective is $\wedge$, a proof has two parts, one for each side. Section 1.2 gives this habit a name, and each later section adds to it.

Quick check. Find the main connective of each statement.

(a) $\neg P \vee Q$

(b) $\neg(P \vee Q)$

(c) $(P \to Q) \wedge (Q \to P)$

Solution.

(a) $\vee$: by convention (1) the $\neg$ applies only to P, so the statement is the disjunction of $\neg P$ and Q.

(b) $\neg$: the parentheses make the whole disjunction the thing denied.

(c) $\wedge$: each parenthesized implication is complete, and the last step joins them with "and".

1.1.9 The form of an argument

Here are two short arguments, one about integers and one about real numbers.

> If n is odd, then n^2 is odd. And n is odd. So n^2 is odd.

> If $x > 2$, then $x^2 > 4$. And $x > 2$. So $x^2 > 4$.

They are about different things, and yet they are the same argument. Write P for "n is odd" and Q for "n^2 is odd" in the first, and P for "$x > 2$" and Q for "$x^2 > 4$" in the second. Both then read: from $P \to Q$ and P, conclude Q. What makes the step correct has nothing to do with odd numbers or squares. It depends only on the shape of the premises, on the fact that one of them is an implication and the other is its hypothesis. That shape is the argument's **form**, and the part of the argument that lives in the letters is its **content**.

Logic studies the forms of arguments, not their content.

Separating the two has a payoff. Once a form has been shown to be correct, every argument of that form is correct, whatever its content; we check the form once and use it forever. An argument form is written with its premises above a line and its conclusion below it:

$$\begin{array}{l} P \to Q \\ P \\ \hline Q \end{array}$$

In running text the same form is written "from $P \to Q$ and P, conclude Q".

1.1.10 Try to break it

Breaking a form means making several statements true at once, and the list of them deserves a name.

Definition 1.1.3. Counterexample Set

The **counterexample set** of an argument form is the list of its premises together with the negation of its conclusion. A **counterexample** to the form is an argument of that form whose premises are **all** true and whose conclusion is false, that is, a choice of statements for the letters that makes **every** member of the counterexample set true. A form is **bad** if it has a counterexample, and **good** if no choice of statements can break it.

For "from $P \to Q$ and Q, conclude P" the counterexample set is $P \to Q$, Q, $\neg P$.

To test a form, try to break it: put statements in for the letters that make every premise true and the conclusion false.

One counterexample settles that the form is bad. Failing to find one settles nothing yet: showing that a form is good takes a reason that covers every choice of statements at once.

Quick check. Write the form of each argument, and try to break it.

(i) If n is divisible by 4, then n is even. And n is even. So n is divisible by 4.

(ii) If $x > 0$ and $y > 0$, then $xy > 0$. And $x > 0$ and $y > 0$. So $xy > 0$.

Solution.

(i) With P: "n is divisible by 4" and Q: "n is even", the form is "from $P \to Q$ and Q, conclude P": the second premise is the conclusion of the implication, not its hypothesis. And $n = 2$ breaks it. The first premise is true for every integer, since a multiple of 4 is $4k = 2(2k)$; the second is true, since 2 is even; yet 2 is not divisible by 4. The form is bad.

(ii) With P the compound statement "$x > 0$ and $y > 0$" (a letter may stand for one) and Q: "$xy > 0$", the form is "from $P \to Q$ and P, conclude Q". It cannot be broken. By the promise table, $P \to Q$ is false only when P is true and Q false; so when $P \to Q$ and P are both true, Q is true. The form is good.

Where to look first. Make the conclusion false first: that fixes some letters, and the premises must then be made true around it. With statements about numbers, try 0, 1, -1, 2 and $\frac{1}{2}$ before anything clever; most bad forms about numbers break at one of them. Take the form "from $P \to Q$ and $Q \to R$, conclude $R \to P$". Its conclusion is false only with R true and P false. Take P: "$x > 2$", Q: "$x > 1$" and R: "$x > 0$"; both premises then hold for every real x. At $x = \frac{1}{2}$, R is true and P is false, so the whole counterexample set $P \to Q$, $Q \to R$, $\neg(R \to P)$ is true, and the form is bad.

Quick check. Write the counterexample set of "from $P \to Q$, conclude Q", and make all of it true with statements about a real number x.

Solution. The counterexample set is $P \to Q$, $\neg Q$. Take P: "$x > 1$" and Q: "$x > 0$". The first member holds for every real x, since a number greater than 1 is greater than 0. At $x = -1$ the second holds too, since $-1 > 0$ is false. So both are true at once, and the form is bad.

Argument B, tried. B's reasoning shows "if n is even, then n^2 is even", and its last sentence concludes the converse, "if n^2 is even, then n is even". That step has the form "from $P \to Q$, conclude $Q \to P$", and it breaks: take P: "n is divisible by 4" and Q: "n is even". The premise is true for every integer, and the conclusion fails at $n = 2$. B's tempting repair, "from $P \to Q$ and Q, conclude P", is the form of (i), which $n = 2$ just broke. With B's own content no integer breaks the repair, since every integer with an even square is even (the fact B set out to prove); but the *form* breaks. A true conclusion reached by a bad form is still not proved.

↻ **Second reading** · *A bad form refutes a step, not its conclusion*

Breaking the last step of Argument B shows that the step proves nothing. It does not show that "if n^2 is even, then n is even" is false, and Argument A proves that it is true. A bad form may even have instances whose premises always give the conclusion. The form "from P, conclude Q" is bad: any true P with a false Q breaks it. Yet in its instance "$x > 2$; so $x > 1$" no number breaks the step, since every real number greater than 2 is greater than 1. The guarantee there comes from a fact about numbers, not from the form. Logic certifies only forms; the content of a step may still be checked by other means.

↻ **Second reading** · *Valid, or only convincing?*

Many true cases are evidence, not a proof. The number n^2+n+41 is prime for $n = 0$, 1, 2, and so on all the way to $n = 39$: forty cases without a failure. At $n = 40$ it is $1600+40+41 = 1681 = 41^2$. An argument from many checked cases can be strong, but it is not of a good form: its premises can all be true while its conclusion is false. Logic books call a good form **valid**, and logic studies only form. That is why this book asks of every step not "is it convincing?" but "could this form be broken?"

A list of checked cases says something about the cases checked, and perhaps about most cases; a proof says something about every case. A scientist who reasons "if the hypothesis is right, we shall observe this; we observe it; so the hypothesis is right" uses the bad form of Quick check (i), "from $P \to Q$ and Q, conclude P" (Section 1.3 names it *affirming the consequent*). The reasoning is a sensible guide to what to test next, and it is not a proof. Some books call an argument of a good form with true premises *sound*; this book keeps that word for a property of rules.

1.1.11 Good form, false sentences

Argument B has true claims and a bad form. The opposite happens too:

If 6 is even, then 6 is prime. And 6 is even. So 6 is prime.

Its form, "from $P \to Q$ and P, conclude Q", is good (Quick check (ii) above), and its conclusion is false. The fault is the first premise, false since 6 is even and not prime. Form and truth are separate questions; all four combinations occur.

	every claim true	some claim false
form good	"If 3 is odd, then 9 is odd. 3 is odd. So 9 is odd."	the argument about 6 above
form bad	Argument B, by its last step "from $P \to Q$, conclude $Q \to P$"	"If 2 is divisible by 4, then 2 is even. 2 is even. So 2 is divisible by 4."

The top left is the odd-square argument at $n = 3$; the bottom right is the Quick check's (i) at $n = 2$. Argument A is not placed. One bad step puts an argument in the bottom row, which is how B got there, but the top row needs every step shown good. A's steps argue from a supposition ("Suppose n is odd ...") and use "some" (an integer k with $n = 2k + 1$), and a form written as premises above a line cannot yet express either. Section 1.4 gives the rules for a supposition, and Chapter 5 those for "some".

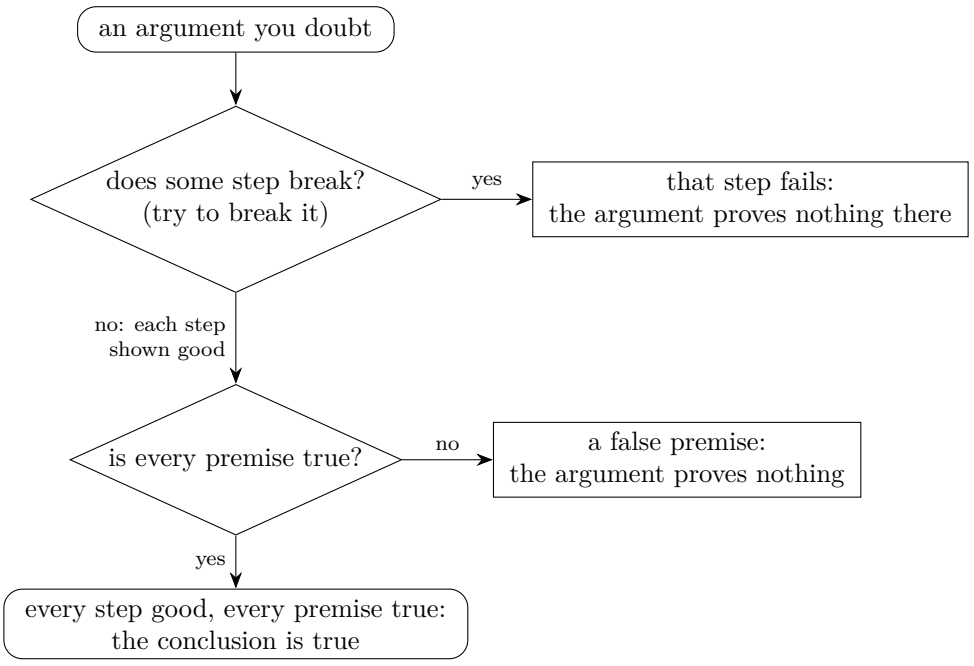
A good form guarantees only that true premises give a true conclusion.

1.1.12 Two ways to attack an argument

An argument can be attacked in two ways, and the two are independent. Attack a **premise**: show that it is false, as with the argument about 6 above. Attack a **step**: show that its form is bad by breaking it, as with Argument B. A false premise says nothing about the steps, and a broken step says nothing about the premises. A defended argument needs both: good forms, and premises the reader accepts.

To doubt an argument, doubt a premise or break a step.

The two attacks, in the order a checker makes them:



Failing to break a step is not the "no" answer: the arrow down from the first question needs every step shown good, by a reason that covers every choice of statements.

1.1.13 Each connective, two rules

We now come to the idea that organizes this chapter and the next. Think about what a mathematician does with the word "and". There are exactly two things to do with it.

- To **prove** a statement of the form $P \wedge Q$, prove P and prove Q.
- To **use** a statement of the form $P \wedge Q$ that is already known, take either half.

Nobody needs to be taught these two moves. They are also, and this is the point, **all** there is to know about "and" for the purpose of writing proofs. Every other fact about $\wedge$, such as that $P \wedge Q$ and $Q \wedge P$ say the same thing, follows from the two moves, as Section 1.2 shows. The first move is called the **introduction rule** for $\wedge$, written $\wedge$I, because it is how a $\wedge$ gets into a proof. The second is the **elimination rule**, $\wedge$E, because it is how a $\wedge$ is taken apart.

The table below is an informal preview. Section 1.2 states these two rules exactly, as patterns of lines.

Rule	What it allows	You write
$\wedge$I	from P and Q, conclude $P \wedge Q$	"We have shown $x > 0$ and $x < 1$. Hence $0 < x < 1$."
$\wedge$E	from $P \wedge Q$, conclude P; also Q	"Since $0 < x < 1$, in particular $x > 0$."

Every connective works the same way. Each comes with introduction rules, saying how to prove a statement with that main connective, and elimination rules, saying what a statement with that main connective lets us conclude. The rules for $\wedge$ come in Section 1.2, those for $\rightarrow$ in Sections 1.3 and 1.4, and those for $\vee$ in Sections 1.5 and 1.6. The rules for $\neg$, $\bot$, and $\leftrightarrow$, and one more rule for proof by contradiction, which stands apart from the others, come after these. Two slogans summarize the plan.

Each connective comes with a way to prove it and a way to use it.

A proof is a chain in which every line is a premise, an assumption, or one application of a rule.

The second slogan is what makes a proof possible to check. Finding a proof may need an idea. Checking one needs none: go through it line by line and ask of each line which rule produced it and whether that rule was applied to lines that really are there. A reader who has lost the thread of a long argument can always fall back on that question, and it is the question to ask of Argument B.

Section 1.4 names exactly what B did wrong; the test *try to break it* has already shown, at $n = 2$ in its Quick check (i), that B's tempting repair, "from $P \rightarrow Q$ and Q, conclude P", is a bad form.

1.1.14 What comes next

Section 1.2 shows what a derivation is, with the rules for "and". Each later section adds one rule. Two questions guide every search: what does the goal's main connective ask for, and what does an available line's main connective let me do?

Remark (Looking ahead). Chapter 5 gives "for every" its rules and explains why an argument about a fixed but arbitrary x proves a statement about every x.

1.1.15 Exercises

1.1.15.1 A. Check your understanding

Exercise. A1

(a) State the definition of a statement. Is "$10^{100}+1$ is prime" a statement? For a real number x, is "$x^2-1=0$" a statement? Is "$0^2-1=0$"?

(b) Let n be an integer, with P: "n is even" and Q: "n is prime". Write "n is prime but not even" in symbols.

(c) Find the main connective of $P \wedge Q \to \neg R$, and write the statement with all its parentheses.

(d) True or false, with a reason: if P and Q are both true, then $P \vee Q$ is false.

(e) Let x and y be real numbers. Write the form of "$x>0$ and $y>0$; so $y>0$ and $x>0$" with premises above a line, saying what each letter stands for.

(f) Which of $\wedge$, $\vee$, $\to$ is the main connective of $(P \vee Q) \wedge (Q \to R)$? Using the preview table of "Each connective, two rules", say what a proof of this statement must do, and name the rule that Section 1.2 will state for it.

Solution.

(a) A statement is a declarative sentence that is either true or false, and not both (Definition 1.1.1). Yes: $10^{100}+1$ is a definite integer, so it either is prime or is not. The definition does not ask us to know which. With x fixed, by the convention of the box in "Statements", "$x^2-1=0$" is a statement about that one number: true if x is 1 or -1, false otherwise. Without the convention it would be an open sentence. The sentence "$0^2-1=0$" is a statement outright, and it is false, since $0^2-1=-1$.

(b) $Q \wedge \neg P$. The word "but" is a conjunction.

(c) Convention (2) groups $P \wedge Q$ before $\to$ acts, and convention (1) lets $\neg$ take only R. So the statement is $(P \wedge Q) \to (\neg R)$, and its main connective is $\to$ (Definition 1.1.2).

(d) False. The mathematical "or" is inclusive, so $P \vee Q$ is true when both parts are true.

(e) With P: "$x>0$" and Q: "$y>0$", the premise is $P \wedge Q$ and the conclusion is $Q \wedge P$:

$$\frac{P \wedge Q}{Q \wedge P}$$

Section 1.2 derives it.

(f) $\wedge$: each parenthesized part is complete, and the last step joins them with "and". So a proof of it has two parts: it proves $P \vee Q$, and it proves $Q \to R$. The rule that puts the two parts together is $\wedge$I. This is the first question of "What comes next", asked before any rule has been stated exactly: the goal's main connective decides the last step.

Exercise. A2: Drill

Answer each part in a line or two.

(a) What is the main connective of $\neg P \wedge Q \to R$?

(b) What is the main connective of $P \vee Q \to R \wedge S$? Write the statement with all its parentheses.

(c) What is the main connective of $\neg(P \to Q) \wedge R$?

(d) Let x be a real number, with P: "$x > 0$" and Q: "$x^2 > x$". Translate "x is positive but x^2 is not greater than x".

(e) Let n be an integer, with O: "n is odd", S: "n^2 is odd" and C: "n^3 is odd". Translate "if n is odd, then n^2 and n^3 are odd".

(f) Let x be a real number, with P: "$x \geq 2$" and Q: "$x \leq 3$". Write "exactly one of $x \geq 2$, $x \leq 3$" in symbols, and decide it at $x = \frac{5}{2}$ and at $x = 4$.

(g) Is "if $x > 1$, then $x^2 > 1$" broken at $x = -2$?

(h) Is "if $x^2 > 1$, then $x > 1$" broken at $x = -2$?

(i) Let n be an integer. Write "n is prime only if n is odd" as an implication, and break it.

(j) Try to break the form "from $P \vee Q$ and P, conclude $\neg Q$".

(k) The argument "If 7 is even, then 7^2 is even. 7 is even. So 7^2 is even." has a false conclusion. Is its form bad?

(l) Take the good form "from $P \to Q$ and P, conclude Q". (i) Give an instance, with statements about numbers, whose conclusion is true although a premise is false. (ii) Can both premises of an instance be false at once?

(m) Give an example, or explain why there is none: an argument of a good form whose premises are all true and whose conclusion is false.

(n) The form "from $P \to Q$ and $Q \to R$, conclude $R \to P$" was broken under "Where to look first". Give an argument of this form about one integer whose premises and conclusion are all true, and say whether it proves its conclusion.

(o) Write the counterexample set of "from $P \to Q$ and $P \vee R$, conclude Q", and make every member true with statements about a real number x.

Solution.

(a) $\to$. The statement is $((\neg P) \wedge Q) \to R$.

(b) $\to$. With all its parentheses it is $(P \vee Q) \to (R \wedge S)$.

(c) $\wedge$. The parentheses keep $\neg$ on $P \to Q$, and the last step joins $\neg(P \to Q)$ and R.

(d) $P \wedge \neg Q$. "But" is a conjunction.

(e) $O \to (S \wedge C)$.

(f) $(P \vee Q) \wedge \neg(P \wedge Q)$. It is false at $\frac{5}{2}$, where both parts hold, and true at 4, where only P holds ($4 \leq 3$ fails).

(g) No. The hypothesis $-2 > 1$ is false, so nothing is promised.

(h) Yes. The hypothesis $4 > 1$ holds and the conclusion $-2 > 1$ fails.

(i) With P: "n is prime" and O: "n is odd", it is $P \to O$. The integer $n = 2$ breaks it: 2 is prime and not odd.

(j) Let x be a real number, with P: "$x \geq 0$" and Q: "$x \leq 0$", and take $x = 0$. Both premises are true (the "or" is inclusive), and $\neg Q$ is false. The form is bad.

(k) No. Its form is "from $P \to Q$ and P, conclude Q", which is good. The premise "7 is even" is false, and a form is bad only when some argument of that form has **all** its premises true and its conclusion false.

(l) (i) "If $2 > 3$, then 3 is odd. $2 > 3$. So 3 is odd." The second premise is false; the first is true, since its hypothesis fails and nothing is promised; and the conclusion is true. A false premise does not make the conclusion false. (ii) No. When P is false, $P \to Q$ is not broken, so it is true; at most one premise of an instance is false.

(m) There is none. Such an argument would be a counterexample to its form, and a good form has none.

(n) "If 8 is divisible by 4, then 8 is even. If 8 is even, then $8 > 0$. So, if $8 > 0$, then 8 is divisible by 4." Every sentence is true, since $8 = 4 \cdot 2$, and the argument proves nothing: its form is bad. The truth of the conclusion is a fact about 8, not a gift of the form. With 6 in place of 8 the premises stay true and the conclusion fails: $6 > 0$, and 4 does not divide 6.

(o) The counterexample set is $P \to Q$, $P \vee R$, $\neg Q$. Take P: "$x > 5$", Q: "$x > 3$", R: "$x < 0$", and $x = -2$. The first member holds for every x; $P \vee R$ holds since $-2 < 0$; and $\neg Q$ holds since $-2 > 3$ is false.

1.1.15.2 B. Practice

Exercise. B1: Translating analysis

Translate each sentence into symbols, using the atomic statements given. You do not need to know what the words mean; only the grammar matters.

(a) "f is continuous at a but not differentiable there", with C: "f is continuous at a" and D: "f is differentiable at a".

(b) "If the sequence (a_n) converges, it is bounded", with V: "(a_n) converges" and B: "(a_n) is bounded".

(c) "x is rational or irrational", with R: "x is rational". (A real number is irrational when it is not rational.)

(d) "The set F is closed if and only if its complement is open", with K: "F is closed" and U: "the complement of F is open".

Solution.

(a) $C \wedge \neg D$. "But" asserts both parts, and "not differentiable there" is the negation of D.

(b) $V \to B$. The pronoun "it" refers to (a_n), so the conclusion is the statement B.

(c) $R \vee \neg R$. "Irrational" means "not rational", so the second part is $\neg R$, not a new atomic statement.

(d) $K \leftrightarrow U$.

Exercise. B2: Parsing

For each statement, insert all the parentheses that the conventions allow us to omit, and name the main connective.

(a) $\neg P \to Q$

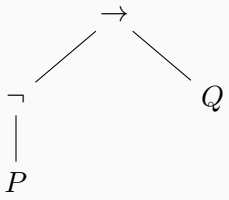
(b) $P \vee \neg Q \leftrightarrow \neg R$

(c) $\neg\neg P \wedge (Q \to R)$

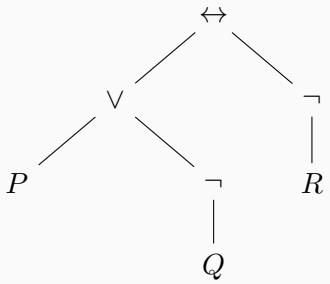
Then let x and y be real numbers, with P: "$x > 0$", Q: "$y > 0$" and R: "$xy > 0$". Write $P \wedge Q \to R$ as an English sentence, and say whether it is true whatever the numbers x and y are.

Solution.

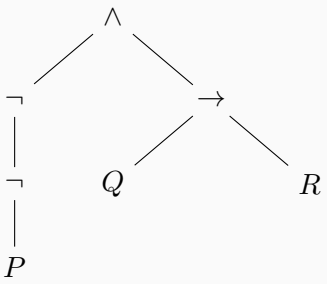
(a) $(\neg P) \to Q$, main connective $\to$. As a tree:



(b) $(P \vee (\neg Q)) \leftrightarrow (\neg R)$, main connective $\leftrightarrow$. Each $\neg$ applies to the single letter after it, and $\vee$ binds more tightly than $\leftrightarrow$. As a tree:



(c) $(\neg(\neg P)) \wedge (Q \to R)$, main connective $\wedge$. The first $\neg$ applies to the statement $\neg P$, which is the smallest complete statement following it. As a tree:



The statement $P \wedge Q \to R$ is $(P \wedge Q) \to R$: "if $x > 0$ and $y > 0$, then $xy > 0$". It is true for all real x and y, since the product of two positive numbers is positive.

Exercise. B3: Necessary and sufficient, used later

Let x be a real number.

(a) Write "$x > 1$ is sufficient for $x^2 > 1$" and "$x > 1$ is necessary for $x^2 > 1$" as implications.

(b) Decide which of the two is true for every real x, and give a value of x that makes the other false.

(c) Find a condition on x, written with $\vee$, that is both necessary and sufficient for $x^2 > 1$. You may use that $x^2 > 1$ exactly when $|x| > 1$, where $|x|$ is x if $x \geq 0$ and $-x$ if $x < 0$.

Solution.

(a) "P is sufficient for Q" means that P is enough to guarantee Q, that is, $P \to Q$. "P is necessary for Q" means that Q cannot hold without P, that is, $Q \to P$. So the two sentences are $(x > 1) \to (x^2 > 1)$ and $(x^2 > 1) \to (x > 1)$.

(b) The first is true for every x: if $x > 1$, then $x > 0$, so $x^2 = x \cdot x > 1 \cdot x = x > 1$, where multiplying $x > 1$ by the positive number x keeps the inequality. The second is false for $x = -2$: then $x^2 = 4 > 1$, but $x < 1$.

(c) $(x > 1) \vee (x < -1)$. It is sufficient: if $x > 1$, then $|x| = x > 1$, and if $x < -1$, then $|x| = -x > 1$; either way $x^2 > 1$. It is necessary: if $x^2 > 1$, then $|x| > 1$, and $|x|$ is x or $-x$, so $x > 1$ or $-x > 1$, the latter meaning $x < -1$.

Exercise. B4: Break it or explain why not

For each form, either find statements about a real number x that make every premise true and the conclusion false, or explain, by checking every case, why no choice of statements can (use the meaning of each connective; for $\to$, the table of "When is an implication broken?").

(a) From $P \to Q$ and $\neg P$, conclude $\neg Q$.

(b) From $P \to Q$ and $Q \to R$, conclude $P \to R$.

(c) From $P \vee Q$, conclude P.

(d) From $P \wedge Q$, conclude $P \vee Q$.

Solution.

(a) Bad. Take P: "$x > 2$" and Q: "$x^2 > 4$", and $x = -3$. The premise $P \to Q$ holds for every x, as the "only if" passage of this section showed. The premise $\neg P$ is true, since $-3 > 2$ is false. And Q is true, since $9 > 4$, so the conclusion $\neg Q$ is false.

(b) It cannot be broken. To make the conclusion $P \to R$ false we need P true and R false. Now look at Q. If Q is true, then $Q \to R$ has a true hypothesis and a false conclusion, so it is broken. If Q is false, then $P \to Q$ has a true hypothesis and a false conclusion, so it is broken. Either way some premise is false.

(c) Bad. Take P: "$x > 0$" and Q: "$x \leq 0$", and $x = -1$. The premise $P \vee Q$ is true, since $-1 \leq 0$, and the conclusion $-1 > 0$ is false.

(d) It cannot be broken. If $P \wedge Q$ is true, then P is true, so $P \vee Q$ is true.

In (b) and (d), "cannot be broken" is shown by reasoning about every case at once, not by failing to find a counterexample: a search that comes up empty settles nothing.

Exercise. B5: One key, sixteen sentences, used later

Let n be an integer. Use only the letters E: "n is even", P: "n is prime", T: "n is a multiple of 3" and G: "$n > 10$". Translate sentences 1–14 into symbols. Then translate 15 and 16 into English, and find an n that makes 16 false.

1. n is even and greater than 10.
2. n is odd.
3. n is an even prime.
4. If n is prime and greater than 10, then n is odd.
5. n is a multiple of 3 but not prime.
6. n is neither even nor a multiple of 3.
7. n is not both even and a multiple of 3.
8. n is prime only if n is odd or at most 10.
9. n is even whenever it is a multiple of 3 greater than 10.
10. Provided $n > 10$, being even is sufficient for n not to be prime.
11. $n > 10$ is necessary for n to be an even multiple of 3.
12. Exactly one of "n is even", "n is prime" holds.
13. At least two of "n is even", "n is prime", "n is a multiple of 3" hold.
14. n is odd, but if it is greater than 10, it is not a multiple of 3.
15. $\neg G \to (E \vee T)$.
16. $(P \wedge \neg E) \to G$.

Solution.

1. $E \wedge G$.
2. $\neg E$: an integer is odd exactly when it is not even.
3. $E \wedge P$.
4. $(P \wedge G) \to \neg E$.
5. $T \wedge \neg P$. "But" is a conjunction.
6. $\neg E \wedge \neg T$.
7. $\neg(E \wedge T)$. Compare 6: one word apart, and different claims (at $n = 2$, 6 is false and 7 is true).
8. $P \to (\neg E \vee \neg G)$. "At most 10" is "not greater than 10", and "P only if Q" is $P \to Q$.
9. $(T \wedge G) \to E$. "P whenever Q" is $Q \to P$.
10. $G \to (E \to \neg P)$. The statement $(G \wedge E) \to \neg P$ says the same; Section 1.4 derives each from the other.
11. $(E \wedge T) \to G$. "Q is necessary for P" is $P \to Q$.
12. $(E \vee P) \wedge \neg(E \wedge P)$.
13. $((E \wedge P) \vee (E \wedge T)) \vee (P \wedge T)$.
14. $\neg E \wedge (G \to \neg T)$.
15. "If n is at most 10, then n is even or a multiple of 3."
16. "If n is an odd prime, then $n > 10$", or "every odd prime is greater than 10". It is false at $n = 3$: 3 is odd and prime, and not greater than 10.

A translation is checked by reading it back, letter by letter, into English. Read 16 back and the reading "odd prime" falls out of $P \wedge \neg E$.

Exercise. B6: Same form?

Here is an argument.

> (T) If 12 is divisible by 6, then 12 is even. If 12 is divisible by 6, then 12 is divisible by 3. So, if 12 is even, then 12 is divisible by 3.

Which of (a)–(d) have the same form as (T)? For one that does, break the form. Is the conclusion of (T) true, and does (T) prove it?

(a) Let x be a real number. If $x > 2$, then $x^2 > 4$. If $x > 2$, then $x > 1$. So, if $x^2 > 4$, then $x > 1$.

(b) Let Δ be a triangle. If Δ is equilateral, then Δ is isosceles. If Δ is equilateral, then each angle of Δ is $60°$. So, if Δ is isosceles, then each angle of Δ is $60°$.

(c) Let n be an integer. If n is divisible by 6, then n is even. If n is even, then n^2 is even. So, if n is divisible by 6, then n^2 is even.

(d) Let x be a real number. If $x > 3$, then $x > 2$. If $x > 3$, then $x > 1$. So, if $x > 3$, then $x > 2$ and $x > 1$.

Solution.

(T) has the form "from $P \to Q$ and $P \to R$, conclude $Q \to R$", with P: "12 is divisible by 6", Q: "12 is even", R: "12 is divisible by 3". The two premises share their hypothesis, and the conclusion joins their two conclusions.

(a) and (b) have this form. In (a), P is "$x > 2$", Q is "$x^2 > 4$" and R is "$x > 1$"; in (b), P is "Δ is equilateral", Q is "Δ is isosceles" and R is "each angle of Δ is $60°$".

(b) has the form "from $P \to Q$ and $Q \to R$, conclude $P \to R$": the second premise starts where the first ends. That is the form of B4(b), which cannot be broken.

(c) has the form "from $P \to Q$ and $P \to R$, conclude $P \to (Q \wedge R)$". It is the look-alike: the same premises as (T), and a different conclusion. It cannot be broken either. To make its conclusion false, P must be true and $Q \wedge R$ false, so Q or R is false; then $P \to Q$ or $P \to R$ has a true hypothesis and a false conclusion, and a premise is false. Section 1.4 derives the forms of (c) and (d).

Breaking the form of (T). Let n be an integer, with P: "n is divisible by 6", Q: "n is even" and R: "n is divisible by 3". Both premises hold for every integer, since $6k = 2 \cdot 3k = 3 \cdot 2k$. At $n = 2$ the conclusion $Q \to R$ is broken: 2 is even and not divisible by 3. So the counterexample set $P \to Q$, $P \to R$, $\neg(Q \to R)$ is true at $n = 2$, and the form is bad. Argument (a) breaks with its own content, at $x = -3$: both premises hold for every x, and $(-3)^2 = 9 > 4$ while $-3 > 1$ fails.

(T) itself. Its conclusion is true, since $12 = 3 \cdot 4$, so an even 12 is divisible by 3. But (T) does not prove it: its form is bad, and the truth of the conclusion comes from a fact about 12, not from the premises.

1.1.15.3 C. Going deeper

Exercise. C1: "Only if" and "unless", used later

Let P and Q be statements.

(a) Explain why "P only if Q" means $P \to Q$, and not $Q \to P$. *(Ask: which situation would make "P only if Q" false?)*

(b) Explain why "P unless Q" means $\neg Q \to P$.

(c) Let x be a real number. Decide whether each sentence is true for every real x. Justify your answer, giving a value of x that makes it false where there is one.
- "$x^2 = 4$ if $x = 2$."
- "$x^2 = 4$ only if $x = 2$."
- "$x^2 > 0$ unless $x = 0$."

Solution.

(a) "P only if Q" says that P does not happen without Q. It is broken exactly when P holds and Q fails. That is also exactly when $P \to Q$ is broken: an implication promises Q whenever P, and the promise fails only when P holds without Q. So the two say the same thing. By contrast, $Q \to P$ is broken when Q holds and P fails, a different situation. For example, "a number is divisible by 4 only if it is even" is true, while "a number is even only if it is divisible by 4" is false (take 2).

(b) "P unless Q" says that P holds in every situation except possibly those where Q holds. So once we know that Q fails, we may conclude P, which is $\neg Q \to P$. For instance, "the equation $ax = 1$ has a solution unless $a = 0$" says: if $a \neq 0$, then $ax = 1$ has a solution.

(c) The first sentence is $(x = 2) \to (x^2 = 4)$. It is true for every x: when $x = 2$, $x^2 = 4$, and when $x \neq 2$ it promises nothing.

The second is $(x^2 = 4) \to (x = 2)$, the same two parts in the opposite direction. It is false for $x = -2$, where $x^2 = 4$ but $x \neq 2$.

The third is $\neg(x = 0) \to (x^2 > 0)$, by (b). It is true for every x: a non-zero real number is positive or negative, and in both cases its square is positive.

The first two sentences differ by one word, and they are implications in opposite directions. Confusing them is exactly the mistake of Argument B at the start of this section.

1.2 Derivations, Starting with And

The previous section stated a slogan: a proof is a chain in which every line is a premise, an assumption, or one application of a rule. This section makes the premises and the rule steps exact (assumptions wait until Section 1.4): what a derivation is, how each line carries its reason, and how anyone can check one. Then it gives the first two rules of logic, the rules for "and".

1.2.1 Why check by matching

A proof in a research paper can run to hundreds of steps, too many to hold in mind at once, so it is checked one step at a time. Two careful readers can disagree about whether a step is "obvious"; a rule that says exactly which steps are allowed settles it, since a step matches the rule or does not. And programs that check proofs work in the same spirit: each step must match a pattern.

Here is a small case: the chain

$$0 < \tfrac{1}{3} < \tfrac{1}{2} < 1$$

makes three claims, $0 < \frac{1}{3}$, $\frac{1}{3} < \frac{1}{2}$, and $\frac{1}{2} < 1$, and a reader who accepts them concludes $0 < 1$. The justification is one fact, used twice:

> if $a < b$ and $b < c$, then $a < c$.

From $0 < \frac{1}{3}$ and $\frac{1}{3} < \frac{1}{2}$ it gives $0 < \frac{1}{2}$. From $0 < \frac{1}{2}$ and $\frac{1}{2} < 1$ it gives $0 < 1$.

To check those two steps, a reader needs nothing about fractions and nothing about what $<$ means. It is enough to see that the two inequalities used in each step have the **same** expression in the middle, and that the new inequality joins the outer two. A clerk who had never seen a number could do it. So a proof can be checked without being understood: a proof is correct when every step matches its pattern, and matching is mechanical.

1.2.2 A rule is a pattern

We borrow one fact about real numbers and use it as a rule.

- **Tr** (for "transitivity", or chaining): from $a < b$ and $b < c$, write $a < c$.

It is a fact about numbers used as a rule, for practice; it is not a rule of logic.

A rule is applied by matching against a small picture of the lines it needs and the line it writes.

A rule is a pattern: given these lines, you may write that line.

Definition 1.2.1. Rule Schema

A **rule schema** is a pattern of lines: **given** lines at the top, possibly none, and exactly **one new** line at the bottom, whose justification names the rule and cites the given lines. Line numbers in a schema are written i, j, k. The letters in the pattern stand for arbitrary **expressions** of the right kind, and a use of the rule replaces them **consistently**: the same letter by the same expression, **symbol for symbol**, everywhere it occurs. A line is **justified by the rule** when it and the lines it cites, with their numbers, are an **instance** of a schema of the rule: the pattern with every letter replaced in this way. The schema of Tr is

i	$a < b$	
j	$b < c$	
k	$a < c$	Tr i, j

Here a, b, c stand for expressions for numbers.

In three points:

- The given lines may stand anywhere above the new line, in any order on the page, and need not be next to each other or to it.
- The citation's order says which line plays which part: "Tr i, j" makes line i play $a < b$ and line j play $b < c$, unless the rule says otherwise ($\wedge$I, below, accepts either order).
- Letters are replaced consistently, symbol for symbol. The middle letter b occurs in both given lines, so it must be replaced by the **same** string of symbols in both, not merely by the same number.

An instance first. Take $i = 3$, $j = 5$, $k = 6$, and replace a by 0, b by $\frac{1}{2}$, and c by 1. The pattern becomes: line 3 is $0 < \frac{1}{2}$, line 5 is $\frac{1}{2} < 1$, and line 6 is $0 < 1$, justified "Tr 3, 5".

Now a noninstance, by a minimal change. Keep line 3 as $0 < \frac{1}{2}$, and let line 5 be $\frac{1}{3} < 1$. Is "6. $0 < 1$, Tr 3, 5" justified? Line 3 forces b to be $\frac{1}{2}$, and line 5 forces it to be $\frac{1}{3}$, so these lines are not an instance of the schema. Line 6 is true, but Tr was not applied, so the line has no justification.

Why define rules this way? Because a rule applied by matching can be checked by someone who knows nothing about the subject. The schema turns "this step is obviously fine" into "this step matches that pattern", a claim anyone can test.

From the rules of logic on, the letters in a schema are Greek: φ, ψ, χ (phi, psi, chi). A Greek letter is a hole in a rule's pattern; a capital P is a statement we are reasoning about. A pattern works like the formula $(a + b)^2 = a^2 + 2ab + b^2$, whose a and b are holes to be filled with any expression, the same one wherever the hole appears. Filling the holes of a schema works the same way, with parentheses put around a compound statement where they are needed. For example, if a schema has the given lines φ and ψ and the new line $\varphi \wedge \psi$, then the three lines $P \to Q$, R, and $(P \to Q) \wedge R$ are an instance of it, with φ replaced by $P \to Q$ and ψ by R.

Warning. The middle expression must match **exactly**, symbol for symbol. The step "$0 < \frac{2}{4}$ and $\frac{1}{2} < 1$, so $0 < 1$" is tempting, because $\frac{2}{4} = \frac{1}{2}$ and every line in it is true, but it is not an application of Tr: the two middle expressions are different strings. A checker who had to notice that they are equal would need arithmetic, and the point of a schema is that a checker needs nothing. Compare symbols, not values.

1.2.3 Derivations without boxes

A proof strings steps together, each using what earlier steps produced. The next definition says exactly how.

Definition 1.2.2. Box-Free Derivation

Fix a list of rules, each given by schemas. A **box-free derivation** is a finite list of numbered lines $1, 2, \ldots, n$, where $n \geq 1$, each carrying one statement and a justification, such that:

(i) the first lines, possibly none and possibly all, are **premises**, justified by nothing; a short horizontal rule is drawn under the last premise;

(ii) every later line is justified by **one** rule of the list, citing **earlier** lines, and it and the cited lines form an instance of a schema of that rule.

The **premises** of the derivation are its premise lines, and its **conclusion** is its last line.

In words. Clause (i) says where a derivation starts: with the statements it is given, which are not proved and need no reason. Clause (ii) says how it goes on: each further line is one step of one rule, and the only lines it may use are ones already written. The word "earlier" is the only condition on which lines may be cited **for now**. Section 1.4 adds assumptions made inside boxes, and then "earlier" is refined to "available": an earlier line may become unusable once the box it sits in has closed.

The layout is always the same. Line numbers run down the left, statements sit in the middle,

and each justification is on the right: the rule's name, then the cited line numbers. A vertical bar runs down the left, and a short horizontal rule under the last premise separates what is given from what is derived. From Section 1.4 on, a box draws a second, shorter bar, and the bars beside a line show which assumptions it depends on.

1.2.4 First derivations

Here are two derivations with Tr, each with the reasoning that found it. The reasoning is the part to watch: a derivation on the page runs from the top down, but it is rarely found that way.

Example. Chaining

Let x, y, u, v, w, z be real numbers. Using the rule Tr only, write a box-free derivation of each conclusion from the given premises.

(a) Premises $0 < x$, $x < y$, $y < 1$; conclusion $0 < 1$.

(b) Premises $u < v$, $w < z$, $v < w$; conclusion $u < z$.

Solution. **(a)** *Idea.* The premises form a chain $0 < x < y < 1$, each link sharing its right end with the next link's left end. So we work **forwards**, from what we have: chain the first two links at their shared middle x, then the result with the third link at y.

1	$0 < x$	
2	$x < y$	
3	$y < 1$	
4	$0 < y$	Tr 1, 2
5	$0 < 1$	Tr 4, 3

Read it back. Lines 1–3 are the premises. Line 4: lines 1 and 2 are an instance of Tr, with a, b, c replaced by 0, x, y. Line 5: lines 4 and 3 are one, with a, b, c replaced by 0, y, 1. The letters are replaced afresh at each use: the b of line 4 has nothing to do with the b of line 5.

(b) *Idea.* This time the chain is scrambled, so we work **backwards**, from what we want. The goal $u < z$ can come from Tr only if we have some $u < m$ and $m < z$ for a single expression m. The only premise starting with u is $u < v$, so try $m = v$; the remaining need is $v < z$, a new, smaller goal. It can come from Tr with some $v < m'$ and $m' < z$; the only premise starting with v is $v < w$, and with $m' = w$ the second need, $w < z$, is a premise. The search has reached the premises, and the derivation is its record, written **forwards**.

1	$u < v$	
2	$w < z$	
3	$v < w$	
4	$v < z$	Tr 3, 2
5	$u < z$	Tr 1, 4

The search could have started instead from $w < z$, the premise ending in z: with $m = w$ it needs $u < w$, which is Tr 1, 3, and then $u < z$ is Tr 4, 2. That is a different correct

derivation. The same premises and conclusion usually have many derivations, and any one of them shows that the conclusion follows from the premises.

Part (b) was found backwards, from the goal; part (a) forwards. The rest of this section names the two directions.

Quick check. Line 1 of a derivation is $0 < x$ and line 3 is $x < 1$. Which is correct: "4. $0 < 1$, Tr $1, 3$" or "4. $0 < 1$, Tr $3, 1$"?

Solution. The first. With the citation "1, 3", line 1 plays $a < b$ and line 3 plays $b < c$: a is 0, b is x in both lines, and c is 1, so the new line $a < c$ is $0 < 1$. With "3, 1", line 3 would play $a < b$ and line 1 would play $b < c$, so b would have to be 1 (from line 3) and 0 (from line 1) at once.

1.2.5 Reading a derivation

Finding a derivation may need an idea, as part (b) of Example (Chaining) did. Checking one never does. For each line after the premises, ask three questions.

1. **Which rule?** Name the rule the justification cites, and check that it is on the list of rules.
2. **Do the cited lines and this line form an instance of that rule's schema?** Find the replacement of the letters: read it off the cited lines in the order the citation gives them (or, for a rule whose lines may be cited in either order, try both: the new line decides), and check that every letter gets the same expression each time it occurs, and that the new line is what the pattern then says.
3. **Are the cited lines earlier?** Each cited number must be smaller than the number of the line doing the citing.

A line that passes all three is correct, whatever it says and whether or not it is true. Here is the check run on part (a) of Example (Chaining).

Line	Rule	Cites	Instance (the letters' values)	Earlier?
1–3	premises	—	—	—
4	Tr	1, 2	a: 0; b: x; c: y	1, 2 < 4: yes
5	Tr	4, 3	a: 0; b: y; c: 1	4, 3 < 5: yes

Every line passes, so the list is a derivation. Section 1.4 extends this table, when "earlier" becomes "available".

Finding a derivation may need an idea; checking one never does.

Second reading · *Two ways to break a derivation*

A line can fail one question and pass the others. The next example breaks a correct derivation in two ways, each time with the smallest change that does it.

Example. Not a derivation

Each list below is the derivation of part (a) of Example (Chaining) with one change. Say which of the three questions of "Reading a derivation" fails, and at which line.

(a) The statements of lines 4 and 5 are swapped, each taking its justification with it, so that line 4 is now "$0 < 1$, Tr $5, 3$" and line 5 is "$0 < y$, Tr $1, 2$".

(b) Line 5 is justified "Tr $3, 4$" instead of "Tr $4, 3$".

Solution.

(a) The list is the following.

1	$0 < x$	
2	$x < y$	
3	$y < 1$	
4	$0 < 1$	Tr 5, 3
5	$0 < y$	Tr 1, 2

Line 5 passes all three questions, exactly as line 4 did before. Line 4 passes the first two: lines 5, 3, and 4 are an instance of Tr with a, b, c replaced by 0, y, 1. It fails the third, because it cites line 5, which is not earlier. So the list is not a derivation. (It would also be useless if it were one: its last line is now $0 < y$, not $0 < 1$.)

(b) Line 5 is now "$0 < 1$, Tr $3, 4$". Lines 3 and 4 are earlier, so the third question passes, and the rule is named, so the first passes. The second fails. The citation makes line 3, $y < 1$, play $a < b$, so a is y and b is 1; and it makes line 4, $0 < y$, play $b < c$, so b is 0. The letter b cannot be both 1 and 0, so there is no instance. The lines are the right ones in the wrong roles.

Change (b) shows why the citation order is part of the justification, and change (a) why the order of the lines is. Exercise C2 below asks what would go wrong if a line were allowed to cite a later one.

1.2.6 What a derivation proves

A derivation shows that its conclusion follows from its premises. The symbol $\vdash$ says that such a derivation exists.

Definition 1.2.3. Derivable

Let Γ be a list of statements $\varphi_1, \ldots, \varphi_n$, where $n \geq 0$, and let ψ be a statement. We say ψ is **derivable** from Γ, and write

$$\Gamma \vdash \psi,$$

if **there is** a derivation, using the rules of logic, whose premises **all belong to** Γ and whose last line is ψ. We write $\varphi_1, \ldots, \varphi_n \vdash \psi$ for $\Gamma \vdash \psi$, and $\vdash \psi$ when Γ is empty.

Here Γ (capital gamma) is just a name for a list of statements.

Definition 1.2.4. Sequent

Let $P_1, \ldots, P_n$ and Q be statements, where $n \geq 0$. The expression $P_1, \ldots, P_n \vdash Q$, read "$P_1$ to P_n prove Q", is a **sequent**. It is the **claim** that Q is derivable from $P_1, \ldots, P_n$ in the sense of Definition 1.2.3. The statements $P_1, \ldots, P_n$ are its **premises** and Q is its **conclusion**; when $n = 0$ we write $\vdash Q$.

The symbol $\vdash$ is called the **turnstile**. A sequent is not true merely because its conclusion is true: it claims that a derivation exists, and it is proved by exhibiting one.

In words, three remarks.

1. *Which derivations?* Until Section 1.4, "derivation" means a box-free derivation (Definition 1.2.2). Section 1.4 widens the word to allow boxes; nothing proved before changes, since a box-free derivation is still a derivation.
2. *Not every premise must be used.* A derivation may ignore some statements of Γ, and it may even have no premise lines at all. It may not use a premise that is **not** in Γ.
3. *Relative to a list of rules.* The symbol $\vdash$ always refers to a list of rules. From "The rules for and" below on, it is the rules of logic introduced so far; adding a rule never destroys a derivation, so a sequent once derived stays derived. Before that, we name the rule: "$\Gamma \vdash \psi$ **with** Tr".

For instance, part (a) of Example (Chaining) shows $0 < x,\ x < y,\ y < 1 \vdash 0 < 1$ with Tr.

In the derivations from here on, the capital letters P, Q, R, S stand for arbitrary statements, as in Section 1.1. A derivation written with them proves not one sequent but infinitely many at once, one for each way of putting statements in for the letters. So a line written with letters is justified only if it is an instance of the rule whatever statements the letters stand for, different statements for different letters included. When we ask whether such a line is an instance, we therefore compare the lines as written, with different letters counted as different symbols: Q is not a conjunct of $(P \wedge Q) \wedge R$, even though it would be one if R happened to be Q.

Quick check. What sequent "with Tr" does part (b) of Example (Chaining) show?

Solution. $u < v,\ w < z,\ v < w \vdash u < z$ with Tr. The premises of the derivation are lines 1–3, and its last line is $u < z$.

Example. Derivations, simplest first

Decide whether each of the following is a derivation, and if so, what it shows.

(a) The single line 1 P premise.

(b) The empty list, with no lines at all.

(c) The derivation of part (a) of Example (Chaining), for the list of rules consisting of Tr alone.

Solution.

(a) Yes. Its only line is a premise, so clause (i) of Definition 1.2.2 holds, and clause (ii) has no lines to speak about. Its last line is P. It shows $P \vdash P$: every statement is derivable from itself. No rule was used, so this holds whatever the list of rules is.

Parts (b) and (c) are answered in the fold below.

Second reading · *Edge cases*

A derivation of one line. The degenerate case is worth a look. A single premise line, with nothing after it, satisfies both clauses of Definition 1.2.2: clause (i) holds, and clause (ii) has no lines to speak about. So it is a derivation, and its premise is its conclusion.

Part (b) of Example (Derivations, simplest first). No. The definition asks for $n \geq 1$ lines, and a derivation must have a last line to be its conclusion. So the empty list derives nothing. In the same way, a list with no premises can never get started with Tr, or any rule that cites a line: its first line would have to cite an earlier line, and there is none. Every rule met before Section 1.4 cites at least one line, so until then $\vdash \psi$ holds for no ψ. Section 1.4 changes that.

Part (c) of Example (Derivations, simplest first). Yes. Lines 1–3 are premises, with the rule drawn under line 3: clause (i). Line 4 cites lines 1 and 2, which are earlier, and lines 1, 2, 4 are an instance of Tr with a, b, c replaced by 0, x, y. Line 5 cites lines 4 and 3, which are earlier, and lines 4, 3, 5 are an instance with a, b, c replaced by 0, y, 1. So clause (ii) holds for both later lines. The premises are the three premise lines and the conclusion is $0 < 1$.

Warning. A premise is **given**, not proved: from the false premises $1 < 0$ and $0 < -1$, the rule Tr gives the false line $1 < -1$, in a correct derivation. What a correct derivation guarantees is only that **if** the premises are true, so is the conclusion, and even that only for rules that never lead from true lines to a false one.

1.2.7 The rules for and

Analysis is full of chains such as $0 < x < 1$, which says $0 < x$ **and** $x < 1$. A proof that **uses** it often needs one half: "in particular $x > 0$, so we may divide by x." A proof that **proves** it must prove both halves; one half is not the whole claim. Section 1.1 previewed these two moves informally. Here they become rules, checked by matching, as Tr was.

To prove "φ and ψ", prove each; to use it, take either.

Definition 1.2.5. Rules for Conjunction

$$\begin{array}{l|l} i & \varphi \\ j & \psi \\ k & \varphi \wedge \psi \end{array} \quad \wedge\text{I } i, j \qquad \begin{array}{l|l} i & \varphi \wedge \psi \\ k & \varphi \end{array} \quad \wedge\text{E } i \qquad \begin{array}{l|l} i & \varphi \wedge \psi \\ k & \psi \end{array} \quad \wedge\text{E } i$$

The first schema is the rule ∧I (**and-introduction**). Its two given lines may be cited in either order: the new line shows which statement is its left conjunct, so question 2 of "Reading a derivation" tries both orders. The second and third are the two forms of the rule ∧E (**and-elimination**), one writing the left conjunct and one the right.

∧I.

- *Schema:* the first one in the block above.
- *Reading:* two lines in hand, anywhere above and cited in either order; write their conjunction.
- *Example:* from line 2, P, and line 5, Q, "∧I 5, 2" (or "2, 5") writes $Q \wedge P$.

∧E.

- *Schema:* the other two in the block.
- *Reading:* one line whose **main connective** (Definition 1.1.2) is ∧; write either side, one layer at a time.
- *Example:* from $(P \wedge Q) \wedge R$ it writes $P \wedge Q$ or R, never P.

If φ and ψ are true, so is $\varphi \wedge \psi$; if $\varphi \wedge \psi$ is true, so is each side.

Warning. ∧E reaches one layer at a time. From line 1, $(P \wedge Q) \wedge R$, the line "Q — ∧E 1" is not an instance of either ∧E schema, because line 1's conjuncts are $P \wedge Q$ and R. The line is wrong even though Q does follow; the repair costs one line: $P \wedge Q$ by ∧E 1 first, then Q from it.

Quick check. Which statements can ∧E write, in one application each, from the line $(P \wedge Q) \wedge (R \wedge S)$?

Solution. Only $P \wedge Q$ and $R \wedge S$. The main connective is the middle ∧, so φ is $P \wedge Q$ and ψ is $R \wedge S$. Each of P, Q, R, S needs a second application, to one of those two lines.

Quick check. Line 1 is $(P \wedge Q) \to R$. May "$P \wedge Q$ — ∧E 1" be written?

Solution. No. Line 1's main connective is →, so ∧E cannot take it apart, although ∧ occurs in it.

1.2.8 Two directions of reasoning

A derivation has premises at the top and its goal, the conclusion we want, at the bottom. Between them is a gap, and finding a derivation means filling the gap. There are two places to start.

From the goal. Ask: **what is the goal's main connective?** If it is $\wedge$, the last line can be written by $\wedge$I, and then the goal $\varphi \wedge \psi$ splits into two **subgoals**, φ and ψ, each to be reached on some line above. This is **goal-directed** reasoning. It works backwards, up from the bottom of the gap, and asks which **introduction** rule builds the goal's main connective.

From the premises. Ask: **what are the main connectives of the lines we already have?** A line whose main connective is $\wedge$ can be taken apart by $\wedge$E, and each conjunct becomes a new line to work with. This is **premise-directed** reasoning. It works forwards, down from the top of the gap, and asks which **elimination** rule uses the line's main connective.

Work from both ends: take the premises apart, and build the goal from its parts.

The derivation is finished when the two ends meet. Part (b) of Example (Chaining) already worked backwards, with Tr in the part that introduction rules play from now on.

Scratch work for a search fits in a two-column table: the **Givens**, what is in hand (premises and lines derived so far), and the **Goal**, what is still needed. A goal-directed step replaces a goal by its subgoals; a premise-directed step adds to the Givens. The search is finished when every goal is among the Givens. The search for $P \wedge Q \vdash Q \wedge P$, one row per step:

Givens	Goal
$P \wedge Q$	$Q \wedge P$
$P \wedge Q$	Q, P ($\wedge$I splits the goal)
$P \wedge Q$, P, Q ($\wedge$E)	Q ✓, P ✓

The derivation this search finds is Proposition 1.2.1 below.

A goal line may carry its rule before its citation is known. When the goal $Q \wedge P$ is written at the bottom and the plan is $\wedge$I, write "$\wedge$I" beside it and leave the line numbers blank, to be filled when the subgoals stand on lines; the left picture in the Idea of Proposition 1.2.1 does this.

1.2.9 First derivations with and

Example. Two warm-ups

Give a derivation of each sequent.

(a) $P,\ Q \vdash Q \wedge P$.
(b) $P \wedge Q \vdash Q$.

Solution.

(a) *Idea.* The goal $Q \wedge P$ has main connective $\wedge$, so, goal-directed, its last line is $\wedge$I with subgoals Q and P. Both subgoals are premises, so the two ends meet at once.

1	P	
2	Q	
3	$Q \wedge P$	$\wedge$I 2, 1

The citation "2, 1" lists Q's line first; "1, 2" would do as well, since the line shows the order.

(b) *Idea.* The goal Q is a letter, with no main connective, so the goal suggests no rule. The premise has main connective $\wedge$, so, premise-directed, take it apart; its right conjunct is the goal.

1	$P \wedge Q$	
2	Q	$\wedge$E 1

Line 2 is the second form of $\wedge$E, with $\varphi := P$ and $\psi := Q$.

The two derivations are box-free derivations in the sense of Definition 1.2.2, with $\wedge$I and $\wedge$E as the list of rules: premises first, and every later line an instance of a schema whose given lines are earlier. By Definition 1.2.3 they show the two sequents. When the goal is a letter, as in (b), only the premises can suggest a step; when every subgoal is a premise, as in (a), the premises need no taking apart. The first proposition below needs both ends.

↻ **Second reading** · *Edge cases for and-introduction*

One line cited twice. Nothing in the schema of $\wedge$I says that i and j are different lines. With $i = j$, the letters φ and ψ are replaced by the same statement, and the rule writes $\varphi \wedge \varphi$. It is a silly line, and a legal one. For instance, $P \vdash P \wedge P$: the goal is a conjunction, so $\wedge$I, with the subgoals P and P, and the one premise meets both.

1	P	
2	$P \wedge P$	$\wedge$I 1, 1

The first result is that the order of a conjunction does not matter.

Proposition 1.2.1. Conjunction Commutes

$P \wedge Q \vdash Q \wedge P$.

Idea. Write the premise at the top, the goal at the bottom, and leave a gap, and write $\wedge$I beside the goal, with its citation still blank (left picture; the lines carry the numbers they will have at the end). Start from the goal. Its main connective is $\wedge$, so the last line will be $\wedge$I, and we need P and Q on earlier lines (right picture: two subgoals, not yet justified). Now start from the premise. Its main connective is $\wedge$, so $\wedge$E gives P and gives Q. That closes the gap: the pieces the goal asks for are the pieces the premise supplies. (This is the search of the Givens/Goal table above.)

1	$P \wedge Q$	
	$\vdots$	
4	$Q \wedge P$	$\wedge$I

1	$P \wedge Q$	
	$\vdots$	
2	P	
3	Q	
4	$Q \wedge P$	$\wedge$I 3, 2

Proof. The following is a box-free derivation with the single premise $P \wedge Q$ and last line $Q \wedge P$.

1	$P \wedge Q$	
2	P	$\wedge$E 1
3	Q	$\wedge$E 1
4	$Q \wedge P$	$\wedge$I 3, 2

Lines 2 and 3 are the two forms of $\wedge$E applied to line 1, and line 4 is $\wedge$I applied to lines 3 and 2, and the line shows Q on the left. By Definition 1.2.3, $P \wedge Q \vdash Q \wedge P$.

■

Read back in words, the derivation is a short paragraph: *We are given that P and Q. In particular P holds. In particular Q holds. So Q and P both hold.* Each sentence is one line, and each "in particular" is one application of $\wedge$E.

Nothing in the derivation looks at what P and Q say: with P: "$0 < x$" and Q: "$x < 1$" it turns "$0 < x$ and $x < 1$" into "$x < 1$ and $0 < x$", and it does the same for any two statements. Because the letters stand for arbitrary statements, one derivation written with letters proves every such instance at once.

The second result is that the grouping of a conjunction does not matter either. It needs the premise taken apart in two layers.

Proposition 1.2.2. Conjunction Is Associative

$(P \wedge Q) \wedge R \vdash P \wedge (Q \wedge R)$.

Idea. From the goal: its main connective is the **outer** $\wedge$, so the last line is $\wedge$I with the subgoals P and $Q \wedge R$, and the second splits again into Q and R. From the premise: $\wedge$E gives $P \wedge Q$ and R, and again, on $P \wedge Q$, gives P and Q. Both searches stop at the three letters, and the ends meet there. Since $\wedge$E reaches one layer at a time, the premise must be peeled in two rounds.

In pictures: the goal below a gap; then the first round of $\wedge$E as lines 2–3, a gap for lines 4–5, and the subgoal $Q \wedge R$ as line 6, lines 6 and 7 not yet justified (line 7 will be $\wedge$I 4, 6).

1	$(P \wedge Q) \wedge R$
	$\vdots$
7	$P \wedge (Q \wedge R)$

1	$(P \wedge Q) \wedge R$	
2	$P \wedge Q$	$\wedge$E 1
3	R	$\wedge$E 1
	$\vdots$	
6	$Q \wedge R$	
7	$P \wedge (Q \wedge R)$	

Proof. The following is a box-free derivation with the single premise $(P \wedge Q) \wedge R$ and last line $P \wedge (Q \wedge R)$.

1	$(P \wedge Q) \wedge R$	
2	$P \wedge Q$	$\wedge$E 1
3	R	$\wedge$E 1
4	P	$\wedge$E 2
5	Q	$\wedge$E 2
6	$Q \wedge R$	$\wedge$I 5, 3
7	$P \wedge (Q \wedge R)$	$\wedge$I 4, 6

Lines 2 and 3 apply $\wedge$E to line 1, with $\varphi := P \wedge Q$ and $\psi := R$. Lines 4 and 5 apply $\wedge$E to line 2, with $\varphi := P$ and $\psi := Q$. Line 6 applies $\wedge$I to lines 5 and 3, and line 7 applies $\wedge$I to lines 4 and 6, with $\varphi := P$ and $\psi := Q \wedge R$. By Definition 1.2.3, $(P \wedge Q) \wedge R \vdash P \wedge (Q \wedge R)$.

■

The lines could come in other orders. Line 3 could come after lines 4 and 5, for instance, since it is used only at line 6. What cannot move is the dependence: every line must come after each line it cites.

Remark. The converse, $P \wedge (Q \wedge R) \vdash (P \wedge Q) \wedge R$, is derived the same way, peeling the premise's inner conjunction on the right instead of the left; Exercise B1 asks you to write it out. With both directions, the brackets in a chain of conjunctions carry no information, which is why the chain $0 < x < 1 < y$ is written without any.

1.2.9.1 A line is not used up

A line of a derivation is a statement that has been established. Citing it does not consume it: it can be cited again, by as many later lines as need it. The derivation of $P \wedge Q,\ R \vdash (P \wedge R) \wedge (Q \wedge R)$ shows this.

Idea. From the goal: $\wedge$I at the bottom, with subgoals $P \wedge R$ and $Q \wedge R$, each split again by $\wedge$I. From the premises: $\wedge$E on line 1 gives P and Q, and R is already a premise. Both subgoals that need R cite the same line 2.

In pictures: the goal below a gap; then the subgoals as lines 5–6, not yet justified.

		1	$P \wedge Q$	
		2	R	
1	$P \wedge Q$	3	P	$\wedge$E 1
2	R	4	Q	$\wedge$E 1
	⋮		⋮	
7	$(P \wedge R) \wedge (Q \wedge R)$	5	$P \wedge R$	
		6	$Q \wedge R$	
		7	$(P \wedge R) \wedge (Q \wedge R)$	$\wedge$I 5, 6

1	$P \wedge Q$		
2	R		
3	P	$\wedge$E 1	
4	Q	$\wedge$E 1	
5	$P \wedge R$	$\wedge$I 3, 2	
6	$Q \wedge R$	$\wedge$I 4, 2	
7	$(P \wedge R) \wedge (Q \wedge R)$	$\wedge$I 5, 6	

Line 2 is cited by lines 5 and 6, and line 1 by lines 3 and 4. In a written proof this is the everyday "we use the hypothesis $x > 0$ again".

↻ **Second reading** · *Any rearrangement of a conjunction*

Take everything apart, rebuild in the new order. One more derivation rearranges four letters: $(P \wedge Q) \wedge (R \wedge S) \vdash (S \wedge Q) \wedge (P \wedge R)$.

Idea. From the premise: two layers of $\wedge$, so $\wedge$E twice on line 1 and then twice on each of the two new lines, which leaves the four letters on lines 4 to 7. From the goal: two layers of $\wedge$, so $\wedge$I builds $S \wedge Q$ and $P \wedge R$ from the letters and then joins them. There is nothing to choose; the only care needed is the order of each citation.

1	$(P \wedge Q) \wedge (R \wedge S)$	
2	$P \wedge Q$	$\wedge$E 1
3	$R \wedge S$	$\wedge$E 1
4	P	$\wedge$E 2
5	Q	$\wedge$E 2
6	R	$\wedge$E 3
7	S	$\wedge$E 3
8	$S \wedge Q$	$\wedge$I 7, 5
9	$P \wedge R$	$\wedge$I 4, 6
10	$(S \wedge Q) \wedge (P \wedge R)$	$\wedge$I 8, 9

The recipe visibly does not depend on these particular letters or brackets. Take the premise apart with $\wedge$E until only letters remain, then build the goal from the letters with $\wedge$I, inside out. It works whenever the premise and the goal are both built from letters by $\wedge$ alone and every letter of the goal occurs in the premise, whatever the bracketing and the order on either side. So "$\wedge$ is commutative and associative" is, in the end, the observation that $\wedge$E can always take a conjunction down to its letters and $\wedge$I can always build one back up.

1.2.10 How this looks in a mathematical proof

Real proofs mix the logical rules with facts about their subject.

Example. A product of numbers in the unit interval

Let x and y be real numbers with $0 < x < 1$ and $0 < y < 1$. Show that $0 < xy < 1$, and name the rule behind each logical step.

Solution. The hypothesis $0 < x < 1$ is the conjunction $(0 < x) \wedge (x < 1)$. By $\wedge$E we have $0 < x$ and $x < 1$ separately, and likewise $0 < y$ and $y < 1$.

Since $x > 0$ and $y > 0$, the product xy is positive. Since $y < 1$ and $x > 0$, multiplying by x gives $xy < x$; together with $x < 1$ this gives $xy < 1$. These steps are school algebra, not logic.

Finally, from $0 < xy$ and $xy < 1$, $\wedge$I gives $(0 < xy) \wedge (xy < 1)$, which is $0 < xy < 1$.

As a derivation, the same proof reads as follows.

1	$(0 < x) \wedge (x < 1)$	
2	$(0 < y) \wedge (y < 1)$	
3	$0 < x$	$\wedge$E 1
4	$x < 1$	$\wedge$E 1
5	$0 < y$	$\wedge$E 2
6	$y < 1$	$\wedge$E 2
7	$0 < xy$	algebra 3, 5
8	$xy < 1$	algebra 3, 4, 6
9	$(0 < xy) \wedge (xy < 1)$	$\wedge$I 7, 8

Lines 7 and 8 are justified "algebra": each is a step by a fact about numbers used as a rule, exactly as Tr was used above, and not a rule of logic. Line 7 uses "a product of two positive numbers is positive"; line 8 uses "if $0 < a$, $a < 1$ and $b < 1$, then $ab < 1$", one fact about numbers (it holds because multiplying $b < 1$ by the positive a gives $ab < a$, and $a < 1$).

The facts about numbers supply the content of the solution; the rules supply the joints, taking statements apart and putting the pieces back together. The goal-directed reading is visible in the prose too: "show $0 < xy < 1$" is a conjunction, so the solution sets itself two subgoals, and it says "finally" when the two meet.

Second reading · *A chain in an analysis proof*

A proof about sequences might contain the line

$$0 < \frac{1}{n+1} < \frac{1}{n} \leq 1 \qquad \text{for } n \in \mathbb{Z}_{>0}.$$

It asserts three links: $0 < \frac{1}{n+1}$, $\frac{1}{n+1} < \frac{1}{n}$, and $\frac{1}{n} \leq 1$. A reader who takes from it a claim about its two ends, $0 < 1$, performs two steps in silence. The first is one application of chaining: $0 < \frac{1}{n+1}$ and $\frac{1}{n+1} < \frac{1}{n}$ give $0 < \frac{1}{n}$. The second joins $0 < \frac{1}{n}$ with $\frac{1}{n} \leq 1$, and it is **not** Tr: the line $\frac{1}{n} \leq 1$ has $\leq$, not $<$, so the lines do not match the schema. It is a different pattern, "from $a < b$ and $b \leq c$, write $a < c$", another true fact about numbers. That is exactly why a justification must name the pattern it uses. "By the chain" does not say which one.

Written chains like this one are everywhere in analysis, and they are derivations with the justifications left off. Nobody writes the justifications out in a real proof about numbers. But a reader checking a hard chain does what the table of "Reading a derivation" does: for each step, which fact, applied to which earlier lines.

1.2.11 Common errors

Besides $\wedge$E applied through a layer, one more error is common.

Warning. ∧E applies only to a line whose **main connective** is ∧. The line $(P \wedge Q) \to R$ contains a ∧, but its main connective is →, so "P — ∧E 1" is illegal. With x a real number, P: "$x > 0$", Q: "$x > 1$" and R: "$x^2 > 1$", the premise "if $x > 0$ and $x > 1$, then $x^2 > 1$" is true for every x, yet at $x = -1$ the line $x > 0$ is false.

1	$(P \wedge Q) \to R$	
2	P	∧E 1

The illegal step leads from a true premise to a false line, so no rule that allowed it could be safe.

1.2.12 The rule card

Before the card, the two rules once more, matched against statements larger than single letters. The schema does not care how big the statements put in for φ and ψ are; it sees only the main connective.

Example. Instances of ∧I and ∧E

Rule	φ	ψ	Cited line(s)	New line
∧I	$P \to Q$	$R \vee S$	$P \to Q$; $R \vee S$	$(P \to Q) \wedge (R \vee S)$
∧E (left)	$\neg P \vee Q$	$Q \to (R \wedge S)$	$(\neg P \vee Q) \wedge (Q \to (R \wedge S))$	$\neg P \vee Q$
∧E (right)	$P \wedge Q$	$(R \to P) \wedge Q$	$(P \wedge Q) \wedge ((R \to P) \wedge Q)$	$(R \to P) \wedge Q$

Not an instance (∧I). From $P \wedge Q$ and R, the line $P \wedge (Q \wedge R)$ is not written by ∧I: the rule joins the two cited lines **whole**, giving $(P \wedge Q) \wedge R$ (or $R \wedge (P \wedge Q)$).

Not an instance (∧E). From $P \wedge (Q \vee R)$, the line Q is not written by ∧E: the right conjunct is $Q \vee R$, and Q sits inside a disjunction, which no rule of this section takes apart.

The card gives each rule of logic so far, its shape, the sentence that uses it in a written proof and, for readers who meet a proof assistant, the Lean code that performs it (`h.1` names the left half of a hypothesis `h`). Nothing later depends on the last column.

Rule	Shape	You write	Lean
∧I	$\varphi, \psi \Longrightarrow \varphi \wedge \psi$	"We show P, and we show Q."	`constructor`
∧E	$\varphi \wedge \psi \Longrightarrow \varphi$ (or ψ)	"In particular, P."	`h.1` (and `h.2`)

Warning (Not allowed: ∧I, ∧E). Each line below is a wrong step, with its reason and the clause of Definition 1.2.2 it breaks.

- "Q from $(P \wedge Q) \wedge R$ by ∧E": reaches through two layers; ∧E writes a conjunct of the cited line's main connective (clause (ii): no instance).

- "P from $(P \wedge Q) \to R$ by $\wedge$E": the main connective is $\to$, not $\wedge$ (clause (ii): no instance).
- "$P \wedge R$ from $P \wedge Q$ and R by $\wedge$I": $\wedge$I joins whole cited lines; P is not a line (clause (ii): no instance).
- "line 4, $Q \wedge P$, $\wedge$I 2, 6": cites line 6, which is not earlier (clause (ii): a later line cited).

Quick check. Line 1 is $(P \vee Q) \wedge (R \wedge S)$. Can "$(R \wedge S) \wedge (P \vee Q)$ — $\wedge$I 1" be written? If not, what must come first?

Solution. No: $\wedge$I cites two lines (or one line twice, which gives a line of the form $\varphi \wedge \varphi$). First write $P \vee Q$ and $R \wedge S$ by $\wedge$E 1; then $\wedge$I joins them in either order.

Remark (Looking ahead). Chapter 4 checks, for every rule of Chapters 1 and 2 at once, that no rule leads from true lines to a false one. Until then each rule comes with the one-sentence reason it is safe, as the sentence after the bullets for "and" did.

1.2.13 Exercises

1.2.13.1 A. Check your understanding

Exercise. A1

(a) State the definition of a box-free derivation and of $\Gamma \vdash \psi$, and write the three schemas of Definition 1.2.5 from memory, naming the rule of each.

(b) Which statements does one application of $\wedge$E give from the line $P \wedge (Q \wedge R)$?

(c) In the derivation in the proof of Proposition 1.2.2, why must line 2 come before line 4?

(d) Fill in the justifications (i)–(iv) of the following derivation.

1	$P \wedge (Q \wedge R)$	
2	P	(i)
3	$Q \wedge R$	(ii)
4	R	(iii)
5	$R \wedge P$	(iv)

(e) True or false, with a reason: if $\Gamma \vdash \psi$ with Tr, and every statement in Γ is true, then ψ is true.

(f) What sequent does the derivation in (d) show?

Solution.

(a) A box-free derivation, relative to a list of rules, is a finite list of numbered lines, at least one, each with a statement and a justification, in which the first lines, possibly none

and possibly all, are premises, justified by nothing, and every later line is justified by one rule of the list, citing earlier lines, so that it and the cited lines form an instance of a schema of that rule (Definition 1.2.2). And $\Gamma \vdash \psi$ means that there is a derivation, using the rules of logic (or a named list of rules, as in "with Tr"), whose premises all belong to Γ and whose last line is ψ (Definition 1.2.3). For the schemas see Definition 1.2.5: $\wedge$I has given lines i: φ and j: ψ and new line k: $\varphi \wedge \psi$, cited "$\wedge$I i, j" (or j, i); the two forms of $\wedge$E each have the given line i: $\varphi \wedge \psi$ and the new line k: φ, respectively ψ, cited "$\wedge$E i".

(b) P, or $Q \wedge R$. The main connective is the first $\wedge$, so φ is P and ψ is $Q \wedge R$.

(c) Line 4 is P, obtained by $\wedge$E from line 2, $P \wedge Q$. A rule may cite only earlier lines (Definition 1.2.2), and P cannot be obtained from line 1 in one step, since line 1's conjuncts are $P \wedge Q$ and R.

(d) (i) $\wedge$E 1; (ii) $\wedge$E 1; (iii) $\wedge$E 3; (iv) $\wedge$I 4, 2. "$\wedge$I 2, 4" is equally right: the line $R \wedge P$ shows which conjunct is on the left.

(e) True. Tr leads from true lines to a true line: if $a < b$ and $b < c$ are true, so is $a < c$. Go down the derivation: the premises are in Γ, so they are true, and each later line comes from earlier, true lines by Tr, so it is true too. Hence the last line, ψ, is true. This is the property the Looking-ahead box of this section promises for the rules of logic.

(f) $P \wedge (Q \wedge R) \vdash R \wedge P$: its only premise is line 1, and its last line is $R \wedge P$.

Exercise. A2: Drill

Answer each part in a line or two.

(a) Lines 1 $2 < y$ and 2 $y < z$ are given. What does "Tr 1, 2" justify?

(b) Line 1 is $((P \wedge Q) \wedge R) \wedge S$. List every statement $\wedge$E can write from line 1 in one step.

(c) Line 3 is $(P \vee Q) \wedge (Q \to R)$. Write the lines that "$\wedge$E 3" can give.

(d) Line 6 is $(R \to S) \wedge P$, justified "$\wedge$I 2, 4", and line 4 is P. What is line 2?

(e) Fill in the justifications of lines 2–6.

1	$(P \wedge Q) \wedge R$	
2	$P \wedge Q$	?
3	R	?
4	Q	?
5	$R \wedge Q$	?
6	$(R \wedge Q) \wedge (P \wedge Q)$	?

(f) Lines 1 $u < 3$ and 2 $4 < v$ are given. Is "$u < v$ — Tr 1, 2" justified?

(g) The premises are 1 $0 < x$, 2 $x < 3$, 3 $3 < y$. Is "4. $0 < y$ — Tr 1, 3" justified?

(h) In the list "1 $0 < 1$, premise; 2 $1 < 2$, premise; 3 $0 < 2$ — Tr 2, 1", which of the three questions of "Reading a derivation" fails?

(i) May $\wedge$E be applied to the line $\neg P \wedge Q$?

(j) Lines 1 $P \wedge (Q \wedge R)$, 2 S and 3 $Q \wedge R$ ($\wedge$E 1) are given. Which of the following can be line 4? (i) $(Q \wedge R) \wedge S$, $\wedge$I 3, 2; (ii) $Q \wedge S$, $\wedge$I 3, 2; (iii) R, $\wedge$E 1; (iv) R, $\wedge$E 3.

(k) Write $0 < \frac{1}{4} < \frac{1}{3} < \frac{1}{2}$ as a derivation of $0 < \frac{1}{2}$ from its three links with Tr.

(l) Derive $P \wedge Q,\ R \vdash R \wedge P$.

(m) Derive $(P \wedge Q) \wedge R \vdash Q$.

(n) Derive $P \wedge Q,\ R \wedge S \vdash S \wedge P$.

(o) Fill in the Givens/Goal table for $P \wedge (Q \wedge R) \vdash R$, one row per step.

Solution.

(a) $2 < z$, with a, b, c replaced by 2, y, z.

(b) $(P \wedge Q) \wedge R$ and S only; $P \wedge Q$, P, Q and R each need further applications.

(c) $P \vee Q$ and $Q \to R$: the main connective is the middle $\wedge$, whatever the conjuncts contain.

(d) $R \to S$. Line 6 has the left conjunct $R \to S$ and the right conjunct P; line 4 supplies P, so line 2 must supply $R \to S$.

(e) 2 $\wedge$E 1; 3 $\wedge$E 1; 4 $\wedge$E 2; 5 $\wedge$I 3, 4 (or 4, 3); 6 $\wedge$I 5, 2. Line 2 is cited twice.

(f) No: the middles 3 and 4 differ, so the lines are no instance of Tr, although $u < v$ does follow (since $3 < 4$).

(g) No: line 1 makes b the expression x, and line 3 makes it 3.

(h) The second: citing line 2 first makes b be 2 (from line 2, as $a < b$) and 0 (from line 1, as $b < c$). The rule is named and the cited lines are earlier.

(i) Yes: $\neg$ binds most tightly, so the main connective is $\wedge$, and $\wedge$E writes $\neg P$ or Q.

(j) (i) Yes. (ii) No: $\wedge$I joins the two cited lines whole, and line 3 is $Q \wedge R$, not Q. (iii) No: R lies a layer deeper in line 1. (iv) Yes.

(k) Premises 1 $0 < \frac{1}{4}$, 2 $\frac{1}{4} < \frac{1}{3}$, 3 $\frac{1}{3} < \frac{1}{2}$; then 4 $0 < \frac{1}{3}$, Tr 1, 2; 5 $0 < \frac{1}{2}$, Tr 4, 3.

(l) 3 P, $\wedge$E 1; 4 $R \wedge P$, $\wedge$I 2, 3.

(m) 2 $P \wedge Q$, $\wedge$E 1; 3 Q, $\wedge$E 2.

(n) 3 P, $\wedge$E 1; 4 S, $\wedge$E 2; 5 $S \wedge P$, $\wedge$I 4, 3.

(o)

Givens	Goal
$P \wedge (Q \wedge R)$	R
$P \wedge (Q \wedge R)$, P, $Q \wedge R$ ($\wedge$E)	R
$P \wedge (Q \wedge R)$, P, $Q \wedge R$, Q, R ($\wedge$E)	R ✓

1.2.13.2 B. Practice

Exercise. B1: The other grouping

Give a derivation of $P \wedge (Q \wedge R) \vdash (P \wedge Q) \wedge R$. For each line, say whether you found it from the goal or from the premise.

Solution. From the goal: $\wedge$I with subgoals $P \wedge Q$ and R, and $P \wedge Q$ by $\wedge$I from P and Q. From the premise: $\wedge$E gives P and $Q \wedge R$, and $\wedge$E on $Q \wedge R$ gives Q and R.

1	$P \wedge (Q \wedge R)$	
2	P	$\wedge$E 1
3	$Q \wedge R$	$\wedge$E 1
4	Q	$\wedge$E 3
5	R	$\wedge$E 3
6	$P \wedge Q$	$\wedge$I 2, 4
7	$(P \wedge Q) \wedge R$	$\wedge$I 6, 5

Lines 2–5 were found from the premise, lines 6–7 from the goal. With Proposition 1.2.2, each of $(P \wedge Q) \wedge R$ and $P \wedge (Q \wedge R)$ is derivable from the other.

Exercise. B2: Annotating a proof

Let x and y be real numbers. Here is a proof.

> Let $1 < x < 2$ and $3 < y < 4$. Then $4 < x + y < 6$.

Write out the proof in full, naming each use of $\wedge$E and $\wedge$I, and marking the steps that are school algebra rather than logic.

Solution. The hypotheses are the conjunctions $(1 < x) \wedge (x < 2)$ and $(3 < y) \wedge (y < 4)$, and the goal is $(4 < x + y) \wedge (x + y < 6)$.

By $\wedge$E, four times, we have $1 < x$, $x < 2$, $3 < y$, and $y < 4$ separately. Adding $1 < x$ and $3 < y$ gives $1 + 3 < x + y$, that is, $4 < x + y$ (school algebra: two inequalities in the same direction may be added). Adding $x < 2$ and $y < 4$ gives $x + y < 6$ (school algebra again). Finally $\wedge$I joins $4 < x + y$ and $x + y < 6$ into $4 < x + y < 6$.

As a derivation it has the same shape as the one in Example (A product of numbers in the unit interval): two premises, four $\wedge$E lines, two lines justified "algebra", and one $\wedge$I line.

Exercise. B3: Two layers, then rebuild

A search for $P \wedge (Q \wedge R),\ S \vdash (R \wedge P) \wedge S$ has reached the stage below. *Name the move, then complete it.*

1	$P \wedge (Q \wedge R)$	
2	S	
	$\vdots$	
7	$(R \wedge P) \wedge S$	?

Solution. The move is *work from both ends* ("Two directions of reasoning"). From the goal: its main connective is $\wedge$, so line 7 is $\wedge$I with the subgoals $R \wedge P$ and S; S is line 2, and $R \wedge P$

splits again, by ∧I, into R and P. From the premise: ∧E in two layers, first P and $Q \wedge R$ from line 1, then R from $Q \wedge R$.

1	$P \wedge (Q \wedge R)$	
2	S	
3	P	∧E 1
4	$Q \wedge R$	∧E 1
5	R	∧E 4
6	$R \wedge P$	∧I 5, 3
7	$(R \wedge P) \wedge S$	∧I 6, 2

The ends meet at line 6: the subgoals R and P are lines 5 and 3, which the premise supplied.

1.2.13.3 C. Going deeper

Exercise. C1: A bad rule

Add to Tr a second rule, Flip: from $a < b$, write $b < a$.

(a) Derive $1 < 0$ from the premise $0 < 1$, using Flip.
(b) Explain why a checker following "Reading a derivation" would accept your derivation.
(c) What property does Tr have that Flip lacks?

Solution.

(a)

1	$0 < 1$	
2	$1 < 0$	Flip 1

(b) Line 1 is a premise. Line 2 names a rule on the list, cites line 1, which is earlier, and lines 1, 2 are an instance of the schema of Flip with a replaced by 0 and b by 1. All three questions pass. The checker never asks whether a line is true, only whether it matches a rule, so a bad rule is invisible to it.

(c) Tr never leads from true lines to a false one (A1(e)). Flip does: from the true line $0 < 1$ it writes the false line $1 < 0$. With such a rule, a derivation from true premises could end anywhere, so checking is only as good as the rules checked against. The rules of logic are chosen to have it; this section's Looking-ahead box says where that is checked.

Exercise. C2: Find the error

A student offers the following list as a derivation of $0 < 4$ with Tr.

1	$0 < 1$	
2	$1 < 2$	
3	$2 < 3$	
4	$3 < 4$	
5	$0 < 3$	Tr 6, 3
6	$0 < 2$	Tr 1, 2
7	$0 < 4$	Tr 5, 4

(a) Is $0 < 1,\ 1 < 2,\ 2 < 3,\ 3 < 4 \vdash 0 < 4$ with Tr?

(b) Is the list a derivation? If not, name the line that fails and the clause of Definition 1.2.2 it breaks.

(c) *(Challenge.)* The two lines "1. $0 < 0$, Tr $2, 2$" and "2. $0 < 0$, Tr $1, 1$" would each pass the first two questions of "Reading a derivation". Use them to explain why a line may cite only earlier lines.

Solution.

(a) Yes. Swap lines 5 and 6 and renumber:

1	$0 < 1$	
2	$1 < 2$	
3	$2 < 3$	
4	$3 < 4$	
5	$0 < 2$	Tr 1, 2
6	$0 < 3$	Tr 5, 3
7	$0 < 4$	Tr 6, 4

Each later line cites earlier lines and is an instance of Tr, with middles 1, 2, 3. So the student's conclusion is derivable (and true).

(b) No. Line 5 cites line 6, which is not earlier, so clause (ii) fails at line 5. Every line is an instance of Tr: lines 6, 3, 5 with middle 2; lines 1, 2, 6 with middle 1; lines 5, 4, 7 with middle 3. So the only fault is the order. That the conclusion is true and derivable does not repair the list: a derivation must meet the definition line by line.

(c) Line 1 cites line 2 twice, and line 2 cites line 1 twice. Each is an instance of Tr with a, b, c all replaced by 0: from $0 < 0$ and $0 < 0$, write $0 < 0$. So if lines could cite later lines, these two lines would justify each other, in a loop that rests on no premise, and the false statement $0 < 0$ would be derivable from nothing. Requiring cited lines to be earlier rules out every such loop: following the citations backwards from any line, the numbers go down, so they must end, and they can end only at premises.

Marking scheme. (10 points) - 3: says the conclusion is derivable and gives a correct derivation (every line an instance, every citation earlier). - 3: names line 5 as the faulty line and the broken clause: a cited line must be earlier. - 1: notes that the other lines are instances of Tr, so the order is the only fault. - 3: explains with the loop why later citations cannot be allowed: the two lines justify each other, rest on no premise, and give a false

line. Fatal errors (0 for the step): calling the list a derivation because its conclusion is true; blaming line 7, which cites line 5, an earlier line.

Exercise. C3: A shortcut through two layers

A student offers the following derivation of $P \wedge (Q \wedge R) \vdash Q \wedge P$.

1	$P \wedge (Q \wedge R)$	
2	Q	$\wedge$E 1
3	P	$\wedge$E 1
4	$Q \wedge P$	$\wedge$I 2, 3

(a) Is line 2 an instance of a schema of $\wedge$E? Justify your answer.

(b) Is the conclusion derivable from the premise? Justify your answer.

(c) The step at line 2 never leads from a true line to a false one. Why must a checker still reject it?

(d) *(Challenge.)* State, as a schema, a rule $\wedge$E* that would allow line 2. Show that every use of $\wedge$E* in a derivation can be replaced by two uses of $\wedge$E.

Solution.

(a) No. To match line 1 with $\varphi \wedge \psi$, we must read φ and ψ off line 1's main connective, which is the first $\wedge$: φ is P and ψ is $Q \wedge R$. So $\wedge$E writes P or $Q \wedge R$, and Q is neither.

(b) Yes. Insert the missing layer.

1	$P \wedge (Q \wedge R)$	
2	P	$\wedge$E 1
3	$Q \wedge R$	$\wedge$E 1
4	Q	$\wedge$E 3
5	$Q \wedge P$	$\wedge$I 4, 2

(c) A derivation is checked by matching each line against the schemas of the rules on the list, and line 2 matches none. A checker that also accepted steps that "obviously follow" would need to understand the statements, and then its verdict would be only as reliable as that understanding. The steps that seem obvious are the ones nobody looks at twice, and they are where unnoticed errors live. (This item is not a find-the-error exercise in the usual sense: the step has no counterexample, since it never leads from truth to falsehood. It is simply not a step of this calculus.)

(d) One such rule, in schema form:

i	$\varphi \wedge (\psi \wedge \chi)$	
k	ψ	$\wedge$E* i

Suppose line k of a derivation is ψ, justified "$\wedge$E* i", with line i the statement $\varphi \wedge (\psi \wedge \chi)$. Replace line k by two lines: first $\psi \wedge \chi$, justified $\wedge$E i (the second form, with the conjuncts φ

and $\psi \wedge \chi$), then ψ, justified by $\wedge$E citing that new line (the first form, with the conjuncts ψ and χ). Renumber the later lines, and every citation of them, up by one; a later citation of line k now cites the second new line, which carries the same statement ψ. Every other line is unchanged and still cites earlier lines. Doing this for each $\wedge$E* line, one at a time, gives a derivation with the same premises and last line that uses only $\wedge$E and $\wedge$I. A rule that can always be replaced in this way adds convenience and no new power; Section 2.3 calls such a rule a derived rule.

Marking scheme. (10 points)

- 2: (a) no, reading φ and ψ off the main connective of line 1.
- 2: (b) yes, with a correct derivation.
- 2: (c) checking is by matching the schemas; a step that is not an instance is rejected, whether or not it is true.
- 2: (d) a correct schema, one given line and one new line.
- 2: (d) the replacement by two $\wedge$E lines, with the renumbering of later citations.

Fatal errors (0 for the step): claiming the conclusion is not derivable; offering a counterexample to line 2.

Exercise. C4: Challenge: What conjunction cannot do

In this exercise read P, Q, R as three different letters, each a symbol with no inner structure. Show that no box-free derivation whose only premise is $P \wedge Q$, and whose only rules are $\wedge$I and $\wedge$E, has a line whose statement contains the letter R. Deduce that, with these two rules, $P \wedge Q \vdash P \wedge R$ cannot be shown.

Solution (Hint 1). Suppose some line does contain R, and look at the **first** such line.

Solution (Hint 2). That first line is not a premise. Which rule wrote it, and which letters can that rule write?

Solution. Suppose, for a contradiction, that some line of such a derivation contains R. The derivation has finitely many lines, so there is a first such line; call it line k. Every line before line k is free of R.

Line k is not a premise, because every premise is $P \wedge Q$, which does not contain R (P and Q being letters other than R). So line k is justified by one of the two rules, citing earlier lines (Definition 1.2.2).

If it is $\wedge$I, citing lines i and j, then its statement is $\varphi \wedge \psi$, where φ and ψ are the statements of lines i and j. Both are earlier than k, so neither contains R, and neither does $\varphi \wedge \psi$.

If it is $\wedge$E, citing line i, then its statement is one conjunct of the statement of line i: a part of it. Line i is earlier than k and free of R, so the part is free of R too.

Either way line k does not contain R, which contradicts the choice of k. So no line contains R. In particular no such derivation has the last line $P \wedge R$, and $P \wedge Q \vdash P \wedge R$ cannot be

shown with these rules.

The hypothesis on the letters matters. If R were the same statement as P, the sequent would be $P \wedge Q \vdash P \wedge P$, which $\wedge$E followed by $\wedge$I 2, 2 derives (as in "One line cited twice"). What the argument shows is that no single derivation works for every choice of R.

The argument, "look at the first bad line; it cannot be a premise, and no rule can create it from good lines", is an induction on the lines of a derivation; Chapter 4 makes such arguments exact.

Marking scheme. (10 points)

- 2: considers a first line containing R, and says why a first one exists (finitely many lines); or, equally, goes down the derivation line by line, showing each line free of R because every line it cites is.
- 2: rules out a premise ($P \wedge Q$ does not contain R, since P, Q, R are different letters).
- 2: $\wedge$I case: the new line is built from the statements of two earlier, R-free lines.
- 2: $\wedge$E case: the new line is part of an earlier, R-free line.
- 2: concludes that no line contains R, and so that $P \wedge R$ is never a last line.

Fatal errors (0 for the step): arguing about one particular derivation instead of every derivation; arguing that R is "not true" instead of about which symbols the rules can write.

1.3 Using an Implication: Modus Ponens

Section 1.2 gave the two rules for "and". This section gives one rule for "implies": the rule that **uses** an implication already in hand. Proving one needs a new device, the subject of Section 1.4; using one is what a proof does every time it applies a theorem. The work is to see exactly what the rule allows, and to find chains of implications forwards and backwards.

1.3.1 Four cards

Before the rule, a puzzle about what "if ... then" says.

Example. Four cards

Each of four cards has a whole number on one side and a color on the other. They lie on a table showing

6, 7, red, green.

Someone claims: *if a card shows an even number on one side, then its other side is red.* Which cards must be turned over to find out whether the claim is true of these four cards? Decide before reading the solution.

Solution. For each card, try to break the claim: can its hidden side make a card that is even and not red? (By the promise reading of Section 1.1, that is the only way to break it.)

The card showing 6. Its number is even, so the hypothesis of the implication holds, and the claim says its other side is red. If that side is green, the claim is false. This card must be turned.

The card showing 7. Its number is odd, so the hypothesis fails, and the claim says nothing

about this card. Whatever its color, the card cannot contradict the claim. Leave it.

The red card. Its color is the **conclusion** of the implication. If its number is even, the card fits the claim; if its number is odd, the card still fits, since the claim says only what even cards look like and nothing about which cards are red. No hidden number can contradict the claim. Leave it.

The green card. If its number is even, it is an even card that is not red, and the claim is false. This card must be turned.

So the answer is the 6 and the green card. Many people choose the 6 and the **red** card. The red card is the tell: choosing it means reading the claim as if it also said "if a card is red, then its number is even", which is the **converse**. An implication and its converse are different claims, and the claim here is only the first.

Two of the cards show the two ways an implication meets a case: the 6 and the red card (not the pair that must be turned). The 6 shows the **hypothesis**, and an implication together with its hypothesis gives its conclusion: that is the rule of this section. The red card shows the **conclusion**, and an implication together with its conclusion gives nothing: that is the most important thing the rule does *not* allow. The green card shows that the conclusion fails; saying exactly what follows from that needs the rules for "not".

1.3.2 The rule

An implication fires on its hypothesis, and gives its conclusion.

Definition 1.3.1. Implication Elimination

i	$\varphi \to \psi$	
j	φ	
k	ψ	$\to$E i, j

This rule is called $\to$E (**implication-elimination**, or **modus ponens**).

We read it in four bullets.

- **Two given lines.** One line is an implication $\varphi \to \psi$, the other is φ. Both stand earlier in the derivation, anywhere above line k and in either order on the page.
- **Read φ and ψ off the main connective.** To see a line as $\varphi \to \psi$, its **main connective** (Definition 1.1.2) must be $\to$. Then φ, the **hypothesis**, is everything that this $\to$ joins on its left, and ψ, the **conclusion**, is everything on its right. In $(P \wedge Q) \to R$, the hypothesis is $P \wedge Q$ and the conclusion is R. In $P \to (Q \to R)$, the hypothesis is P and the conclusion is $Q \to R$, itself an implication.
- **The hypothesis must match exactly.** Line j must be the statement φ, symbol for symbol: the letter φ occurs in both given lines, and a use of the rule replaces it by the same statement in both (Definition 1.2.1). A part of φ will not do, and neither will a statement that "says the same thing". From $(P \wedge Q) \to R$ and $Q \wedge P$, the rule gives nothing, although $Q \wedge P$ and $P \wedge Q$ are each derivable from the other (Proposition 1.2.1); the hypothesis must first be rebuilt as $P \wedge Q$, by $\wedge$E and $\wedge$I.
- **The new line is the whole conclusion.** The rule writes ψ, the whole of what stands

right of the main $\to$, and nothing else. If ψ is $Q \wedge R$, the rule writes $Q \wedge R$, and getting Q takes one more line, by $\wedge$E.

The citation lists the implication first and its hypothesis second, always: "$\to$E i, j", as in the schema. Only ψ can come out, so the order changes nothing; it makes every $\to$E line read the same way.

↻ **Second reading** · *A legal, useless line*

The statements φ and ψ may be the same. From $P \to P$ and P, the rule writes P, which was already there. The line is legal and useless.

Why this rule. It is what "by Theorem so-and-so" means. For integers d and n, write $d \mid n$, read "d divides n", when $n = dk$ for some integer k. A theorem of school arithmetic says: if an integer is divisible by 2 and by 3, then it is divisible by 6. For $n = 42$ it says

$$\big((2 \mid 42) \wedge (3 \mid 42)\big) \to (6 \mid 42).$$

We know $2 \mid 42$ ($42 = 2 \cdot 21$) and $3 \mid 42$ ($42 = 3 \cdot 14$). $\wedge$I builds the hypothesis, and $\to$E fires: $6 \mid 42$, which is right, as $42 = 6 \cdot 7$. As a derivation, with the theorem's instance and the two facts as premises:

1	$((2 \mid 42) \wedge (3 \mid 42)) \to (6 \mid 42)$	
2	$2 \mid 42$	
3	$3 \mid 42$	
4	$(2 \mid 42) \wedge (3 \mid 42)$	$\wedge$I 2, 3
5	$6 \mid 42$	$\to$E 1, 4

Every application of a theorem has this shape: check the theorem's hypotheses, assemble them into exactly the theorem's hypothesis, and fire.

Quick check. To get "36 divides 72" from the theorem "if 4 and 9 divide n, then 36 divides n" at $n = 72$, which line, besides the theorem's instance itself, must stand on the page before $\to$E fires?

Solution. The conjunction $(4 \mid 72) \wedge (9 \mid 72)$, exactly: the theorem's hypothesis at $n = 72$. It is built by $\wedge$I from the lines $4 \mid 72$ and $9 \mid 72$, which hold because $72 = 4 \cdot 18$ and $72 = 9 \cdot 8$. Then $\to$E gives $36 \mid 72$, and indeed $72 = 36 \cdot 2$.

Why it is safe. If $\varphi \to \psi$ and φ are true, then ψ is true: that is what the promise says.

Why nothing more. From $\varphi \to \psi$ alone, $\to$E cannot fire: an implication by itself asserts neither its hypothesis nor its conclusion. And $\to$E does not run backwards: from $\varphi \to \psi$ together with its conclusion ψ, it gives nothing about φ. That pattern,

$$\varphi \to \psi,\ \psi \ \text{ therefore } \ \varphi,$$

is called **affirming the consequent**, and it is the red card of the puzzle. It is not an instance of $\to$E, because the second given line of $\to$E must be the hypothesis φ, and here it is the conclusion ψ. It is a bad form: Section 1.1 broke it at $n = 2$. It is also the step that would repair Argument B.

Warning. Affirming the consequent looks like modus ponens and is not. Let x be a real number. "If $x > 2$, then $x^2 > 4$" is true for every x, and at $x = -3$ we have $x^2 = 9 > 4$; but $x > 2$ is false there. From an implication and its **conclusion**, nothing follows about the hypothesis.

Quick check. Line 1 is $P \to (Q \to R)$. What does $\to$E give from line 1 and a line Q? And from line 1 and a line P?

Solution. From line 1 and Q: nothing. The main connective of line 1 is the first $\to$, so its hypothesis is P, and Q does not match it. From line 1 and P: the conclusion $Q \to R$. A second application, to $Q \to R$ and the line Q, would then give R.

The hypothesis can itself be an implication, and then the brackets put it on the left.

Quick check. Line 1 is $(P \to Q) \to (R \wedge P)$ and line 2 is $P \to Q$. What does $\to$E 1, 2 give? And what does $\to$E give from line 1 and a line P?

Solution. $R \wedge P$. The main connective of line 1 is the $\to$ outside the brackets, so its hypothesis is $P \to Q$, and line 2 is that whole hypothesis. From line 1 and P: nothing. P is only a part of the hypothesis $P \to Q$.

Remark. At this point an implication can be **used** but not **proved**. A derivation can write an implication only by taking it out of a line that already contains it: $\wedge$E applied to $(P \to Q) \wedge R$, or $\to$E applied to $P \to (Q \to R)$ and P. No rule so far builds a new $\to$. Section 1.4 gives the rule that does.

↻ **Second reading** · *A rule is not one more premise*

Picture a reader who grants both "if $x > 2$, then $x^2 > 4$" and "$x > 2$", yet will not write "$x^2 > 4$". We might try to persuade this reader by adding the step itself as a third premise. With P: "$x > 2$" and Q: "$x^2 > 4$", the new premise is

$$((P \to Q) \wedge P) \to Q,$$

"if (if $x > 2$, then $x^2 > 4$) and $x > 2$, then $x^2 > 4$". Nothing is gained. Getting from the three premises to Q still takes two steps of the rules: $\wedge$I to join the first two premises into one line, and $\to$E to fire the new premise on that line. A reader who would not take one step may refuse

these two in the same way. A fourth premise that states this passage would call for a fifth, and so on without end: a step that is written down as a premise never gets taken.

So →E is not a statement standing among the premises. It is a **permission** to write a line, and that is why a derivation names its rules apart from its premises (Definition 1.2.2). The puzzle is an old one: it was first told, as a dialogue, in 1895.

1.3.3 Chains of implications

With →E, a derivation can pass along a chain: from P and $P \to Q$ to Q, from Q and $Q \to R$ to R, and so on. Such a chain can be found from either end.

Forwards, from the premises. An available implication (for now, one on an earlier line) whose hypothesis is also available can **fire**: →E writes its conclusion, which becomes available in turn and may be the hypothesis of another implication. Keep firing until the goal appears. This is premise-directed reasoning, as in Section 1.2.

Backwards, from the goal. Suppose the goal is ψ. Look among the available lines for an implication whose **conclusion** is ψ, say $\varphi \to \psi$. If there is one, it is enough to reach φ: then →E writes ψ. So φ becomes the new goal, and we repeat. This is goal-directed reasoning of a new kind: in Section 1.2 the goal's main connective chose an introduction rule, but here the goal may be a letter, and it still tells us which implication to look for.

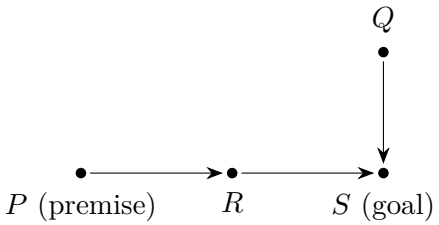
The two directions meet in the middle, when a line reached forwards is a goal reached backwards. In written mathematics the backward step has a standard phrase, "to show ψ, it suffices to show φ".

In the Givens/Goal table, →E fired forwards adds ψ to the Givens; worked backwards, when $\varphi \to \psi$ is among the Givens, it replaces the goal ψ by φ.

Quick check. The premises are P, $P \to R$, $Q \to S$, and $R \to S$, and the goal is S. Search backwards: which implications could give the goal, and which of them leads to a derivation?

Solution. Two premises have conclusion S: $Q \to S$ and $R \to S$. The first makes Q the new goal, and no line is Q or has conclusion Q, so that branch is a dead end. The second makes R the new goal; $P \to R$ has conclusion R, and its hypothesis P is a premise. So the derivation is R by →E from $P \to R$ and P, then S by →E from $R \to S$ and R. A backward search may branch, and a branch that stalls is abandoned for the next one.

A picture shows the search at a glance. Draw each statement as a dot and each implication as an arrow from its hypothesis to its conclusion. Forwards, follow the arrows out of what you have; backwards, follow the arrows into the goal.



Backwards from S, two arrows arrive. One comes from Q, which no arrow reaches and no premise gives: a dead end. The other comes from R, which P reaches.

Example. Chains of implications

Give a derivation of each sequent.

(a) $P \to Q,\ P \vdash Q$.
(b) $P,\ P \to Q,\ Q \to R \vdash R$.
(c) $P \wedge Q,\ P \to (Q \to R) \vdash R$.
(d) $(P \wedge Q) \to R,\ P,\ Q \vdash R$.
(e) $(P \to Q) \wedge (R \to S),\ P \wedge R \vdash Q \wedge S$.

Solution.

(a) *Idea.* There is nothing to choose: the premises are an implication and its hypothesis, and the rule fires once.

1	$P \to Q$	
2	P	
3	Q	$\to$E 1, 2

This is the form of the two short arguments at the start of "The form of an argument" in Section 1.1 (n odd, so n^2 odd; $x > 2$, so $x^2 > 4$), now shown as a sequent: a derivation with these premises and this last line exists.

(b) *Idea, forwards.* The premise $P \to Q$ has its hypothesis P available, so it fires and gives Q. Now $Q \to R$ has its hypothesis available, so it fires and gives R, the goal.

1	P	
2	$P \to Q$	
3	$Q \to R$	
4	Q	$\to$E 2, 1
5	R	$\to$E 3, 4

Idea, backwards. The goal R is a letter. Which available implication has conclusion R? Only $Q \to R$, so it suffices to reach Q (left picture: the goal at the bottom, its new subgoal Q above it). Which implication has conclusion Q? Only $P \to Q$, so it suffices to reach P; and P is a premise, so the gap closes (right picture).

1	P	
2	$P \to Q$	
3	$Q \to R$	
	⋮	
4	Q	
5	R	→E 3, 4

1	P	
2	$P \to Q$	
3	$Q \to R$	
4	Q	→E 2, 1
5	R	→E 3, 4

Both searches write the same five lines, the forward one from the top and the backward one from the last line up. Exercise B3 has a case where the backward search is much shorter.

(c) *Idea.* The goal R is a letter. From the goal: the only line that mentions R is $P \to (Q \to R)$, but its conclusion is $Q \to R$, not R; the R sits one implication deeper. From the premises: $P \wedge Q$ is a conjunction, so take it apart by ∧E, giving P and Q. Now P is exactly the hypothesis of $P \to (Q \to R)$, which fires and gives $Q \to R$. This new line is an implication whose hypothesis Q is available, so it fires too. In the pictures, the goal stands alone at first; then ∧E writes lines 3 and 4, and only line 5 is missing; line 6 stays unjustified until line 5 is there for →E to cite.

1	$P \wedge Q$	
2	$P \to (Q \to R)$	
	⋮	
6	R	

1	$P \wedge Q$	
2	$P \to (Q \to R)$	
3	P	∧E 1
4	Q	∧E 1
	⋮	
6	R	

1	$P \wedge Q$	
2	$P \to (Q \to R)$	
3	P	∧E 1
4	Q	∧E 1
5	$Q \to R$	→E 2, 3
6	R	→E 5, 4

Line 6 is the first line in this chapter where →E fires on an implication that was itself **written by** →E, not given as a premise.

(d) *Idea.* This time the goal leads. The goal R is the conclusion of $(P \wedge Q) \to R$, so it suffices to reach its hypothesis $P \wedge Q$. That subgoal is a conjunction, so, goal-directed as in Section 1.2, it comes from ∧I with subgoals P and Q, and both are premises. In the pictures, the goal stands alone at first; then the subgoal $P \wedge Q$, read off line 1, sits above it as line 4, still unjustified.

1	$(P \wedge Q) \to R$	
2	P	
3	Q	
	$\vdots$	
5	R	

1	$(P \wedge Q) \to R$	
2	P	
3	Q	
	$\vdots$	
4	$P \wedge Q$	
5	R	$\to$E 1, 4

1	$(P \wedge Q) \to R$	
2	P	
3	Q	
4	$P \wedge Q$	$\wedge$I 2, 3
5	R	$\to$E 1, 4

Compare (c). There the implication took its two hypotheses one at a time, $P \to (Q \to R)$, and the conjunction was a premise to be taken apart. Here the implication takes both hypotheses at once, $(P \wedge Q) \to R$, and the conjunction is a subgoal to be built. The content is the same; the brackets decide which rules the derivation uses. (Section 1.4 shows that each of the two implications is derivable from the other.)

(e) *Idea.* Both ends at once. From the goal: $Q \wedge S$ is a conjunction, so its last line is $\wedge$I, with subgoals Q and S. From the premises: both are conjunctions, so $\wedge$E takes each apart, giving $P \to Q$, $R \to S$, P, and R. The two ends meet through $\to$E: $P \to Q$ fires on P and gives the subgoal Q, and $R \to S$ fires on R and gives the subgoal S. In the second picture lines 3–6 came from the premises and line 9 from the goal; lines 7–8, still unjustified, are where the two searches meet.

1	$(P \to Q) \wedge (R \to S)$	
2	$P \wedge R$	
	$\vdots$	
9	$Q \wedge S$	

1	$(P \to Q) \wedge (R \to S)$	
2	$P \wedge R$	
3	$P \to Q$	$\wedge$E 1
4	$R \to S$	$\wedge$E 1
5	P	$\wedge$E 2
6	R	$\wedge$E 2
	$\vdots$	
7	Q	
8	S	
9	$Q \wedge S$	$\wedge$I 7, 8

1	$(P \to Q) \wedge (R \to S)$	
2	$P \wedge R$	
3	$P \to Q$	$\wedge$E 1
4	$R \to S$	$\wedge$E 1
5	P	$\wedge$E 2
6	R	$\wedge$E 2
7	Q	$\to$E 3, 5
8	S	$\to$E 4, 6
9	$Q \wedge S$	$\wedge$I 7, 8

Each derivation in the example is a box-free derivation (Definition 1.2.2) with $\wedge$I, $\wedge$E, and $\to$E as its rules, so by Definition 1.2.3 each shows its sequent.

1.3.4 Checking a line of modus ponens

Finding a chain can take some searching; checking one never does. For a line justified "$\to$E i, j", the questions of Section 1.2 (which rule? is it an instance? are the cited lines earlier?) become four small ones:

1. Are lines i and j earlier than this line?
2. Is the main connective of line i a $\to$? If so, read off φ (its left side) and ψ (its right side).
3. Is line j exactly φ, symbol for symbol?
4. Is this line exactly ψ?

Here is the check run on lines 5 and 6 of (c), and on a tempting wrong line that might replace them, "R — $\to$E 2, 4".

Line	Cites	φ, ψ read off line i	Line j is φ?	Line is ψ?
5: $Q \to R$	2, 3	P, $Q \to R$	yes (P)	yes
6: R	5, 4	Q, R	yes (Q)	yes
wrong: R	2, 4	P, $Q \to R$	no (Q is not P)	no (R is not $Q \to R$)

The wrong line fails twice: it feeds line 2 the wrong hypothesis and takes out a part of the conclusion. The main connective settles it: the first $\to$ is the one applied last, so everything after it, $Q \to R$, is the conclusion.

1.3.5 Two chains about numbers

The chains so far were made of letters. Here are two made of facts about numbers, one found forwards and one backwards. In both, the premises are true school facts, and the work is to see which implication fires on what.

Example. Two chains about numbers

(a) Let n be an integer. From the five premises

$$6 \mid n, \qquad (6 \mid n) \to ((2 \mid n) \wedge (3 \mid n)), \qquad (2 \mid n) \to (4 \mid n^2),$$
$$(3 \mid n) \to (9 \mid n^2), \qquad ((4 \mid n^2) \wedge (9 \mid n^2)) \to (36 \mid n^2),$$

derive $36 \mid n^2$.

(b) Let x and y be real numbers. From the six premises

$$x > 1, \qquad y > 1, \qquad (x > 1) \to (x^2 > 1), \qquad (y > 1) \to (y^2 > 1),$$
$$((x^2 > 1) \wedge (y^2 > 1)) \to (x^2 + y^2 > 2), \qquad (x > 2) \to (x^2 + y^2 > 2),$$

derive $x^2 + y^2 > 2$.

Solution.

(a) *The premises.* Each implication is a fact about every integer. If $n = 6k$, then $n = 2(3k)$ and $n = 3(2k)$. If $n = 2k$, then $n^2 = 4k^2$; if $n = 3k$, then $n^2 = 9k^2$. And a number divisible by 4 and by 9 is divisible by 36, the theorem of the Quick check on 72.
Idea, forwards. Fire every implication whose hypothesis is in hand. Only the second premise can fire at first, on $6 \mid n$. Its conclusion is a conjunction, which $\wedge$E splits into $2 \mid n$ and $3 \mid n$. Now the third and fourth premises fire. The last premise wants its whole hypothesis on one line, so $\wedge$I assembles it exactly, and the last premise fires. In the pictures, the goal stands alone at first; then lines 6–8 are written forwards, and line 11, still unjustified, is the hypothesis the last premise asks for.

1	$6 \mid n$
2	$(6 \mid n) \to ((2 \mid n) \wedge (3 \mid n))$
3	$(2 \mid n) \to (4 \mid n^2)$
4	$(3 \mid n) \to (9 \mid n^2)$
5	$((4 \mid n^2) \wedge (9 \mid n^2)) \to (36 \mid n^2)$
	$\vdots$
12	$36 \mid n^2$

Then lines 6–8, written forwards:

1	$6 \mid n$	
2	$(6 \mid n) \to ((2 \mid n) \wedge (3 \mid n))$	
3	$(2 \mid n) \to (4 \mid n^2)$	
4	$(3 \mid n) \to (9 \mid n^2)$	
5	$((4 \mid n^2) \wedge (9 \mid n^2)) \to (36 \mid n^2)$	
6	$(2 \mid n) \wedge (3 \mid n)$	$\to$E 2, 1
7	$2 \mid n$	$\wedge$E 6
8	$3 \mid n$	$\wedge$E 6
	$\vdots$	
11	$(4 \mid n^2) \wedge (9 \mid n^2)$	
12	$36 \mid n^2$	$\to$E 5, 11

The finished derivation:

1	$6 \mid n$	
2	$(6 \mid n) \to ((2 \mid n) \wedge (3 \mid n))$	
3	$(2 \mid n) \to (4 \mid n^2)$	
4	$(3 \mid n) \to (9 \mid n^2)$	
5	$((4 \mid n^2) \wedge (9 \mid n^2)) \to (36 \mid n^2)$	
6	$(2 \mid n) \wedge (3 \mid n)$	$\to$E 2, 1
7	$2 \mid n$	$\wedge$E 6
8	$3 \mid n$	$\wedge$E 6
9	$4 \mid n^2$	$\to$E 3, 7
10	$9 \mid n^2$	$\to$E 4, 8
11	$(4 \mid n^2) \wedge (9 \mid n^2)$	$\wedge$I 9, 10
12	$36 \mid n^2$	$\to$E 5, 11

Read back, the twelve lines say: if 6 divides n, then 36 divides n^2.

(b) *The premises.* Each is true for all real x and y. If $x > 1$, then $x^2 = x \cdot x > x > 1$, and the same for y. If $x^2 > 1$ and $y^2 > 1$, adding gives $x^2 + y^2 > 2$. If $x > 2$, then $x^2 > 4$, and $y^2 \geq 0$, so $x^2 + y^2 > 4 > 2$.

Idea, backwards. Two premises end in the goal. The last one makes $x > 2$ the subgoal, but no line is $x > 2$ and no implication ends in it: a dead end, abandoned. The fifth premise makes its hypothesis, a conjunction, the subgoal. $\wedge$I splits it into $x^2 > 1$ and $y^2 > 1$, and each of these is the conclusion of an implication whose hypothesis is a premise. In the pictures, the goal stands alone at first; then the subgoal read off premise 5 sits above it as line 9, still unjustified.

1	$x > 1$
2	$y > 1$
3	$(x > 1) \to (x^2 > 1)$
4	$(y > 1) \to (y^2 > 1)$
5	$((x^2 > 1) \wedge (y^2 > 1)) \to (x^2 + y^2 > 2)$
6	$(x > 2) \to (x^2 + y^2 > 2)$
	$\vdots$
10	$x^2 + y^2 > 2$

Then the subgoal read off premise 5:

1	$x > 1$	
2	$y > 1$	
3	$(x > 1) \to (x^2 > 1)$	
4	$(y > 1) \to (y^2 > 1)$	
5	$((x^2 > 1) \wedge (y^2 > 1)) \to (x^2 + y^2 > 2)$	
6	$(x > 2) \to (x^2 + y^2 > 2)$	
	$\vdots$	
9	$(x^2 > 1) \wedge (y^2 > 1)$	
10	$x^2 + y^2 > 2$	$\to$E 5, 9

The finished derivation:

1	$x > 1$	
2	$y > 1$	
3	$(x > 1) \to (x^2 > 1)$	
4	$(y > 1) \to (y^2 > 1)$	
5	$((x^2 > 1) \wedge (y^2 > 1)) \to (x^2 + y^2 > 2)$	
6	$(x > 2) \to (x^2 + y^2 > 2)$	
7	$x^2 > 1$	$\to$E 3, 1
8	$y^2 > 1$	$\to$E 4, 2
9	$(x^2 > 1) \wedge (y^2 > 1)$	$\wedge$I 7, 8
10	$x^2 + y^2 > 2$	$\to$E 5, 9

Premise 6 is true and never cited. A true premise can still be a dead end: what decides a backward search is whether the subgoal can be reached, not whether the implication is true.

Quick check. In (b), why is premise 6 never fired?

Solution. Its hypothesis $x > 2$ stands on no line, and no line has it as its conclusion, so →E has nothing to fire it on.

1.3.6 How this looks in a mathematical proof

Let x be a real number with $x > 2$. Suppose we know two facts of school algebra:

- if $x > 2$, then $x^2 > 4$ (since $x > 2 > 0$, multiplying $x > 2$ by x gives $x^2 > 2x$, and $2x > 4$);
- if $x^2 > 4$, then $x^2 + 1 > 5$ (add 1 to both sides).

A written proof that $x^2 + 1 > 5$ takes one sentence: *Since $x > 2$, we have $x^2 > 4$, and so $x^2 + 1 > 5$.* As a derivation, with the two facts and the hypothesis as premises, it reads as follows.

1	$(x > 2) \to (x^2 > 4)$	
2	$(x^2 > 4) \to (x^2 + 1 > 5)$	
3	$x > 2$	
4	$x^2 > 4$	→E 1, 3
5	$x^2 + 1 > 5$	→E 2, 4

Each piece of the sentence is one line.

Line	Sentence
3	"Since $x > 2$, . . . " (the hypothesis)
4	". . . we have $x^2 > 4$, . . . " (→E, firing the first fact)
5	". . . and so $x^2 + 1 > 5$." (→E, firing the second fact)

The sentence never mentions lines 1 and 2: writers leave a known implication unstated. Each "since φ, ψ" in a proof is one application of →E to an implication the reader is trusted to supply. When a proof is hard to follow, writing those implications down is often how one finds the gap: an implication is false, or its hypothesis was never established.

1.3.7 Common errors

The first warning of this section was about the wrong given line: the conclusion in place of the hypothesis. Another common error is a hypothesis that is only partly there.

Warning. →E needs the **whole** hypothesis. From $(P \wedge Q) \to R$ and P alone, it gives nothing. Take the true theorem "if $x > 0$ and $y > 0$, then $xy > 0$" and the line $x > 0$: at $x = 1$, $y = -1$ the theorem and $x > 0$ are true, and $xy > 0$ is false.

Second reading · *The same slip in analysis*

This is the commonest slip in analysis: "f is continuous, so by the theorem f attains a maximum". The theorem meant has a second hypothesis, that f is defined on a closed bounded interval, and without it the conclusion fails: $f(x) = x$ on the open interval $(0, 1)$ is continuous and has no largest value, since $\frac{a+1}{2}$ is a larger value than a for every a in $(0, 1)$.

A near relative is a hypothesis in the wrong order, as in the third bullet of the rule's reading: $Q \wedge P$ is not the hypothesis of $(P \wedge Q) \to R$. The repair is two $\wedge$E lines and one $\wedge$I line, and Exercise B1 asks for it.

Warning. A third error: firing an arrow that is not the whole line. $\to$E fires only on a **whole line** whose main connective is $\to$. From $P \wedge (Q \to R)$ and Q, first take out $Q \to R$ by $\wedge$E; the step straight to R is not an instance of the rule, although it never leads from truth to falsehood. From $(P \to Q) \vee R$ and P, $\to$E cannot fire at all, and the line Q does not follow. Let x be a real number, with P: "$x > 2$", Q: "$x > 4$", R: "$x \leq 4$". The first line holds for every x: if $x \leq 4$, then R is true, and if $x > 4$, then Q is true, so $P \to Q$ is. At $x = 3$ the line P is true and Q is false.

1.3.8 The rule card

Before the card, the rule at work on large statements. Each row is one use of $\to$E: the two given lines, and the line it writes.

Example. Instances of $\to$E

φ	ψ	The implication	The hypothesis line	New line
$P \wedge (Q \vee R)$	$S \to P$	$(P \wedge (Q \vee R)) \to (S \to P)$	$P \wedge (Q \vee R)$	$S \to P$
$P \to Q$	$R \wedge (S \vee P)$	$(P \to Q) \to (R \wedge (S \vee P))$	$P \to Q$	$R \wedge (S \vee P)$
$(P \vee Q) \wedge \neg R$	$(P \wedge S) \vee Q$	$((P \vee Q) \wedge \neg R) \to ((P \wedge S) \vee Q)$	$(P \vee Q) \wedge \neg R$	$(P \wedge S) \vee Q$

Not an instance. From $(P \to Q) \to R$ and P, $\to$E writes neither Q nor R: the hypothesis is $P \to Q$, and P is only a part of it.

The card of Section 1.2 grows by one row. *Shape* shows the given lines, then the new line; *You write* is the phrase a written proof uses for the rule.

Rule	Shape	You write	Lean
$\wedge$I	$\varphi, \psi \Longrightarrow \varphi \wedge \psi$	"We show P, and we show Q."	`constructor`
$\wedge$E	$\varphi \wedge \psi \Longrightarrow \varphi$ (or ψ)	"In particular, P."	`h.1` (and `h.2`)

Rule	Shape	You write	Lean
→E	$\varphi \to \psi,\ \varphi \Longrightarrow \psi$	"Since P implies Q, and P holds, Q."	`exact h hp`

In the Lean column, `h` is the implication and `hp` the proof of its hypothesis; `h hp` applies the one to the other, which is modus ponens.

Warning (Not allowed: →E). Each step below is wrong, for the reason after the colon; the reason names the bullet of the rule's reading that it breaks.

- "R from $P \to (Q \to R)$ and Q": Q is not the hypothesis, which is P (*the hypothesis must match exactly*).
- "P from $P \to Q$ and Q": the second given line must be the hypothesis, not the conclusion; this is affirming the consequent (*two given lines*).
- "R from $(P \wedge Q) \to R$ and P": the whole hypothesis must stand on a line (*the hypothesis must match exactly*).
- "Q from $P \to (Q \wedge R)$ and P": the new line is the whole conclusion, $Q \wedge R$ (*the new line is the whole conclusion*).
- "R from $P \wedge (Q \to R)$ and Q": the cited line's main connective is $\wedge$; take out $Q \to R$ first (*read φ and ψ off the main connective*).

Remark (Looking ahead). Chapter 4 proves more than "affirming the consequent is a bad form": no derivation, however long, works for every choice of P and Q and leads from $P \to Q$ and Q to P.

1.3.9 Exercises

1.3.9.1 A. Check your understanding

Exercise. A1

(a) Write the schema of →E from memory.

(b) Lines 1 and 2 of a derivation are $P \to (Q \wedge R)$ and P. Which of the following can line 3 be, by one application of →E? $Q \wedge R$; Q; $R \wedge Q$; $P \to Q$.

(c) In Example (Four cards), why is the red card not worth turning?

(d) True or false, with a reason: from lines $P \to Q$ and $Q \to R$, →E gives $P \to R$.

Solution.

(a) Given lines i: $\varphi \to \psi$ and j: φ; new line k: ψ, justified "→E i, j" (Definition 1.3.1).

(b) Only $Q \wedge R$, the conclusion of line 1. The line Q needs a further ∧E, and $R \wedge Q$ two further ∧E lines and a ∧I line. The line $P \to Q$ is not the conclusion of line 1, so →E cannot write it.

(c) The claim is "even → red", and the red card shows the conclusion. An odd number behind it would not contradict the claim, which says nothing about which cards are

red; an even number would fit it. Inferring anything about its number from its color would be affirming the consequent.

(d) False. →E needs an implication and a line that is its hypothesis. The hypothesis of $Q \to R$ is Q, and $P \to Q$ is not Q; the hypothesis of $P \to Q$ is P, which is not a line. So →E does not apply at all. (The sequent $P \to Q,\ Q \to R \vdash P \to R$ does hold, but its derivation needs the rule of Section 1.4.)

Exercise. A2: Drill

Each part takes a line or two. Lines are numbered as in a derivation. The parts run from one step, through a missing line, a justification column and steps to judge, to short derivations.

(a) Lines 1 $(P \vee Q) \to R$, 2 $P \vee Q$. What does "→E 1, 2" give?

(b) Lines 1 $P \to Q \wedge R$, 2 P. What does "→E 1, 2" give?

(c) Lines 1 $((P \vee Q) \wedge R) \to (S \to P)$, 2 $(P \vee Q) \wedge R$. What does "→E 1, 2" give?

(d) Line 7 is $R \wedge S$, justified "→E 4, 6", and line 6 is $P \to Q$. What is line 4?

(e) Line 5 is Q, justified "→E 2, 3", and line 2 is $(R \wedge S) \to Q$. What is line 3?

(f) Justify each line, the first two being premises: 1 $(P \to Q) \wedge R$, 2 $R \to P$, 3 $P \to Q$, 4 R, 5 P, 6 Q.

(g) Lines 1 $P \to (Q \to R)$, 2 Q. Is "→E 1, 2" legal?

(h) Lines 1 $(P \wedge Q) \to R$, 2 $Q \wedge P$. Is "R — →E 1, 2" legal?

(i) Lines 1 $P \to Q$, 2 R, 3 $R \wedge P$. Is "4. Q — →E 1, 3" legal?

(j) Lines 1 $(P \to Q) \wedge R$, 2 P. Is "3. Q — →E 1, 2" legal? If not, repair it.

(k) Break the form of "if n is divisible by 6, then by 3; n is divisible by 3; so n is divisible by 6".

(l) Premises $P \to Q$, $R \to Q$, R; goal Q. Searching backwards, which implication leads to a derivation?

(m) Derive $P \wedge (P \to Q) \vdash Q$.

(n) Derive $P \to (Q \to R),\ P,\ Q \vdash R$.

(o) Derive $P \to (Q \wedge R),\ (Q \wedge R) \to S,\ P \vdash S \wedge Q$.

Solution.

(a) R, the conclusion of line 1, whose hypothesis $P \vee Q$ is line 2 exactly.

(b) $Q \wedge R$: since $\wedge$ binds more tightly, line 1 is $P \to (Q \wedge R)$. Getting Q needs one more ∧E line.

(c) $S \to P$: the hypothesis of line 1 is $(P \vee Q) \wedge R$, which is line 2 exactly, and its conclusion is the implication $S \to P$.

(d) $(P \to Q) \to (R \wedge S)$. →E cites the implication first, so line 4 is an implication whose hypothesis is line 6 and whose conclusion is line 7.

(e) $R \wedge S$, the hypothesis of line 2, exactly.

(f) Lines 1 and 2 are premises; 3 is ∧E 1; 4 is ∧E 1; 5 is →E 2, 4; 6 is →E 3, 5. An implication cut out of a conjunction fires on a line that →E wrote.

(g) No: the hypothesis of line 1 is P, and line 2 is Q.

(h) No: the hypothesis must be $P \wedge Q$ exactly, and $Q \wedge P$ is a different statement.

(i) No: line 3 is $R \wedge P$, not the hypothesis P. First write P by $\wedge$E 3 as line 4, then Q by $\to$E 1, 4.

(j) No: the main connective of line 1 is $\wedge$, so line 1 is not an implication, and $\to$E cannot fire on it. Repair: 3 $P \to Q$, $\wedge$E 1; 4 Q, $\to$E 3, 2.

(k) Take $n = 3$. The first premise is true for every integer, since $6k = 3 \cdot 2k$; the second is true, since $3 = 3 \cdot 1$; and 3 is not divisible by 6. Both premises true, conclusion false: the form is affirming the consequent.

(l) $R \to Q$: its hypothesis R is a premise, so $\to$E fires at once. The other, $P \to Q$, makes P the new goal, and nothing gives P: a dead end.

(m) 2 P, $\wedge$E 1; 3 $P \to Q$, $\wedge$E 1; 4 Q, $\to$E 3, 2.

(n) 4 $Q \to R$, $\to$E 1, 2; 5 R, $\to$E 4, 3.

(o) 4 $Q \wedge R$, $\to$E 1, 3; 5 S, $\to$E 2, 4; 6 Q, $\wedge$E 4; 7 $S \wedge Q$, $\wedge$I 5, 6.

1.3.9.2 B. Practice

Exercise. B1: Two chains

Give a derivation of each sequent. For each, say which step you found first, and whether it came from the goal or from a premise.

(a) $P,\ P \to Q,\ Q \to (R \wedge S) \vdash S \wedge Q$.

(b) $(P \wedge Q) \to R,\ Q \wedge P \vdash R$.

Solution.

(a) The first step found was from the goal: $S \wedge Q$ is a conjunction, so the last line is $\wedge$I with subgoals S and Q. Then from the premises: $P \to Q$ fires on P and gives Q, and $Q \to (R \wedge S)$ fires on Q and gives $R \wedge S$, which $\wedge$E takes apart.

1	P	
2	$P \to Q$	
3	$Q \to (R \wedge S)$	
4	Q	$\to$E 2, 1
5	$R \wedge S$	$\to$E 3, 4
6	S	$\wedge$E 5
7	$S \wedge Q$	$\wedge$I 6, 4

Line 4 is cited twice, by line 5 and by line 7.

(b) The first step found was from the goal: R is the conclusion of line 1, so it suffices to reach $P \wedge Q$ exactly. The premise $Q \wedge P$ is not it. So take the premise apart and rebuild in the right order.

1	$(P \wedge Q) \to R$	
2	$Q \wedge P$	
3	Q	$\wedge$E 2
4	P	$\wedge$E 2
5	$P \wedge Q$	$\wedge$I 4, 3
6	R	$\to$E 1, 5

Exercise. B2: Two facts and a number

Let x be a real number. Take as premises the two facts "if $0 < x < 1$, then $x^2 < x$" and "if $x^2 < x$, then $x^3 < x^2$", and the hypothesis $0 < x < 1$. Derive $x^3 < x^2$, naming the rule on each line. Then write the argument as one or two sentences of prose.

Solution. The hypothesis $0 < x < 1$ is the conjunction $(0 < x) \wedge (x < 1)$, which is exactly the hypothesis of the first fact, so the first fact fires on it at once.

1	$((0 < x) \wedge (x < 1)) \to (x^2 < x)$	
2	$(x^2 < x) \to (x^3 < x^2)$	
3	$(0 < x) \wedge (x < 1)$	
4	$x^2 < x$	$\to$E 1, 3
5	$x^3 < x^2$	$\to$E 2, 4

In prose: *Since* $0 < x < 1$*, we have* $x^2 < x$*, and therefore* $x^3 < x^2$. Both facts are true for every real x: if $0 < x < 1$, multiplying $x < 1$ by $x > 0$ gives $x^2 < x$; and if $x^2 < x$, then $x(x-1) < 0$, so $x > 0$, and multiplying $x^2 < x$ by x gives $x^3 < x^2$.

Exercise. B3: Which premises are needed?

Find a derivation of

$$P,\ P \to Q,\ P \to R,\ R \to S,\ S \to U,\ Q \to T \ \vdash\ T,$$

and list the premises it uses. Explain why a backward search finds the needed premises faster than a forward one here.

Solution. Search backwards. The goal T is the conclusion of $Q \to T$ only, so it suffices to reach Q. The goal Q is the conclusion of $P \to Q$ only, so it suffices to reach P, a premise.

1	P	
2	$P \to Q$	
3	$P \to R$	
4	$R \to S$	
5	$S \to U$	
6	$Q \to T$	
7	Q	$\to$E 2, 1
8	T	$\to$E 6, 7

It uses premises 1, 2, and 6. The others are never cited, which is allowed: a derivation need not use every premise (Definition 1.2.3). Deleting lines 3–5 and renumbering gives a derivation of $P,\ P \to Q,\ Q \to T \vdash T$.

A forward search fires every implication whose hypothesis is available, relevant or not: from P it writes Q and R, then S from R, then U from S, and T from Q — five new lines, three of them useless. The backward search looks only at implications whose conclusion is the current goal, and at each stage exactly one qualifies, so it writes only the two lines the derivation needs.

Exercise. B4: Assemble the hypothesis

A search for $P \to (Q \wedge R),\ (R \wedge P) \to S,\ P \vdash S$ has reached the stage below. *Name the move, then complete it.*

1	$P \to (Q \wedge R)$	
2	$(R \wedge P) \to S$	
3	P	
	$\vdots$	
7	S	?

Solution. The move is to work **backwards from the goal**. The goal S is the conclusion of line 2, so it suffices to reach line 2's hypothesis $R \wedge P$, exactly, and then fire. That subgoal is a conjunction: P is line 3, and R comes from line 1, which fires on P and gives $Q \wedge R$, whose right half is R.

1	$P \to (Q \wedge R)$	
2	$(R \wedge P) \to S$	
3	P	
4	$Q \wedge R$	$\to$E 1, 3
5	R	$\wedge$E 4
6	$R \wedge P$	$\wedge$I 5, 3
7	S	$\to$E 2, 6

Line 6 writes $R \wedge P$, with R on the left, because that is how the hypothesis of line 2 is written; $P \wedge R$ would not match it. Line 3 is cited twice, by lines 4 and 6.

Exercise. B5: Supply the missing premise

In (a)–(c), add **one** premise, not the conclusion itself, so that the sequent can be derived with $\wedge$I, $\wedge$E and $\to$E, and give the derivation. In (d), each sentence about an integer n leaves an implication unsaid: write it, and decide whether it is true for every integer n.

(a) $P \to Q,\ Q \to R \vdash R$.

(b) $P,\ (P \wedge Q) \to R \vdash R$.

(c) $P,\ Q \vdash R \wedge P$.

(d) (i) "Since 4 divides n, n is even." (ii) "Since n is even, n^2 is divisible by 4." (iii) "Since n^2 is even, n is even." (iv) "Since n^2 is divisible by 4, n is divisible by 4."

Solution.

(a) Several premises work. Add P:

1	$P \to Q$	
2	$Q \to R$	
3	P	
4	Q	$\to$E 1, 3
5	R	$\to$E 2, 4

Or add Q: then 4 R, $\to$E 2, 3. Any premise from which these rules reach P or Q (for instance $P \wedge Q$) is right too.

(b) Several premises work. Add Q:

1	P	
2	$(P \wedge Q) \to R$	
3	Q	
4	$P \wedge Q$	$\wedge$I 1, 3
5	R	$\to$E 2, 4

Or add $P \wedge Q$ itself (then one $\to$E line), or $Q \wedge P$ (then two $\wedge$E lines and one $\wedge$I line

first). Any premise from which these rules reach Q or $P \wedge Q$ is right.

(c) Several premises work: any premise from which, together with P and Q, these rules reach R. For instance $P \to R$:

1	P	
2	Q	
3	$P \to R$	
4	R	$\to$E 3, 1
5	$R \wedge P$	$\wedge$I 4, 1

$Q \to R$ works the same way, and so does $(P \wedge Q) \to R$, with an $\wedge$I line first. The premise R alone also works, and it says the most: it hands over the new part of the conclusion outright.

(d) (i) "If $4 \mid n$, then n is even": true, since $4k = 2(2k)$. (ii) "If n is even, then $4 \mid n^2$": true, since $n = 2k$ gives $n^2 = 4k^2$. (iii) "If n^2 is even, then n is even": true, by Argument A of Section 1.1. (iv) "If $4 \mid n^2$, then $4 \mid n$": false at $n = 2$, where $4 \mid 4$ and 4 does not divide 2. The sentence in (iv) looks just like the other three. A missing premise is where a hidden error hides.

1.3.9.3 C. Going deeper

Exercise. C1: Find the error

Here is an argument, followed by the derivation its author offers for it.

> If 10 is divisible by 6, then 10 is even. And 10 is even. So 10 is divisible by 6.

With D: "10 is divisible by 6" and E: "10 is even":

1	$D \to E$	
2	E	
3	D	$\to$E 1, 2

(a) Name the line that fails, and say why it is not an instance of the schema of $\to$E.

(b) Show that the step can lead from true premises to a false conclusion.

(c) What new statement about divisibility can be correctly derived from the two premises with the rules so far?

Solution.

(a) Line 3. The schema needs, besides the implication $\varphi \to \psi$, a line that is its hypothesis φ. Matching line 1 gives $\varphi := D$ and $\psi := E$, so the second cited line would have to be D; line 2 is E, which is ψ. The line affirms the consequent.

(b) Try to break the step: this instance already does. The first premise is true, because every multiple of 6 is even ($6k = 2(3k)$); the second is true, since $10 = 2 \cdot 5$; and the conclusion is false: $6k$ is at most 6 when $k \leq 1$ and at least 12 when $k \geq 2$, so $10 = 6k$ for no integer k. So the step at line 3 leads from two true premises to a false line, and no safe rule can allow it.

(c) None. Go down any derivation from these premises line by line, as in Exercise (C4: Challenge: What conjunction cannot do): every line is $D \to E$, E, or a conjunction built from them by $\wedge$ alone. A premise is of this kind. $\wedge$I joins two such lines into another one. $\wedge$E gives back a conjunct of such a line, which is again of this kind. $\to$E would need, besides the implication $D \to E$, a line D, and D is not of this kind; so $\to$E never fires, and no line is D. What we can reach are the premises and conjunctions of them, such as $(D \to E) \wedge E$, which say nothing the premises did not. The premises tell us nothing about whether 10 is divisible by 6, which is why the conclusion had to be smuggled in.

Exercise. C2: Both ends

Consider the sequent

$$P \to (Q \to R),\ S \to P,\ S \wedge Q \ \vdash\ R \wedge S.$$

(a) Describe a search for a derivation in words, in three parts: what the goal asks for; what the premises give; where the two meet.

(b) Write the derivation, naming the rule on every line.

Solution (Hint 1). Two premises have main connective $\to$ and one has $\wedge$; the goal has $\wedge$. Write down what each main connective lets you do.

Solution (Hint 2). The goal splits into R and S. Which available implication could end in R, and after how many firings?

Solution.

(a) *From the goal:* $R \wedge S$ is a conjunction, so the last line is $\wedge$I with subgoals R and S. The only line mentioning R is $P \to (Q \to R)$, whose conclusion is $Q \to R$; so R needs $Q \to R$ and Q, and $Q \to R$ needs P. *From the premises:* $S \wedge Q$ splits by $\wedge$E into S and Q; then $S \to P$ fires on S and gives P. *Where they meet:* S is at once a subgoal and a piece of the premise; P, reached forwards, is the hypothesis the backward search asked for; and Q, from $\wedge$E, is the last hypothesis R needs.

(b) The derivation, with the rule on every line:

1	$P \to (Q \to R)$	
2	$S \to P$	
3	$S \wedge Q$	
4	S	$\wedge$E 3
5	Q	$\wedge$E 3
6	P	$\to$E 2, 4
7	$Q \to R$	$\to$E 1, 6
8	R	$\to$E 7, 5
9	$R \wedge S$	$\wedge$I 8, 4

Line 4 is cited twice, by line 6 and by line 9.

Marking scheme. (10 points)

- 2: from the goal, splits $R \wedge S$ into the subgoals R and S for a final $\wedge$I.
- 2: takes $S \wedge Q$ apart by $\wedge$E.
- 3: fires $S \to P$, then $P \to (Q \to R)$, then $Q \to R$, each $\to$E citing an implication and exactly its hypothesis.
- 2: the final $\wedge$I, whose line is $R \wedge S$.
- 1: the three-part description names where the searches meet.

Any other correct derivation earns the same points step for step (its lines may come in another order, as long as each cites only earlier lines). Fatal errors (0 for the step): $\to$E on $P \to (Q \to R)$ with Q (the wrong hypothesis); affirming the consequent anywhere.

1.4 Proving an Implication: the Rule of the Box

Sections 1.2 and 1.3 gave three rules. $\wedge$I builds a conjunction; $\wedge$E and $\to$E take a line apart or use it. With them we can **use** an implication, but we cannot **prove** one. This section fills that gap with the rule $\to$I. The rule asks us to do something new: to assume a statement we do not know, inside a boundary that says where the assumption holds. The boundary is a **box**. With the box we can name the error of Argument B, and say exactly which earlier lines a line may cite.

1.4.1 Supposing something we do not know

Here is a small theorem about a real number x.

If $x > 2$, then $x^2 > 4$.

The proof everyone would write goes like this. *Suppose $x > 2$. Then $x > 0$, so multiplying $x > 2$ by x gives $x^2 > 2x$. Also $2x > 4$. Hence $x^2 > 4$.* Then, often without saying it: *this proves that if $x > 2$, then $x^2 > 4$.*

Look at the first sentence. We do not know that $x > 2$. The number x is arbitrary, and it may

well be 0. So "suppose $x > 2$" asserts nothing. It opens a **temporary world**, one in which $x > 2$ holds, and asks what else holds there.

Inside that world, the sentences "$x^2 > 2x$", "$2x > 4$", and "$x^2 > 4$" are true. Outside it they are not: at $x = 0$ all three are false. Only the last sentence of the proof leaves the temporary world. What it carries out is not "$x^2 > 4$" but the implication "if $x > 2$, then $x^2 > 4$", and that is true for every real x.

None of the rules so far can do this: ∧I writes a conjunction, and ∧E and →E write a piece of a line that is already there. To prove "if $x > 2$, then $x^2 > 4$" from nothing, we must be allowed to suppose.

A supposition is only safe if it has a boundary. Without one, the proof above could go on: "... Hence $x^2 > 4$. So $x^2 > 4$ for every real x." That is false at $x = 0$. The error is not in any computation; it is that a sentence from inside the temporary world was used outside it.

A derivation draws the boundary as a **box**. The assumption opens the box: it is written with a vertical bar to its left and a short horizontal rule under it. Every line derived under the assumption carries the same bar. When the box closes, the bar stops, and exactly one thing leaves: the implication from the assumption to the box's last line.

To prove "if φ, then ψ", assume φ in a box, reach ψ inside it, and close the box.

1.4.2 The introduction rule for implication

From now on the given part of a schema may also be a **box**, cited as a range of lines from its first line to its last; nothing else in Definition 1.2.1 changes.

Here is the rule, as a pattern. Beside it, for reference, is the rule →E of Definition 1.3.1, so that the two rules for "implies" stand in one place.

Definition 1.4.1. Rules for Implication

i	φ	asm, for →I
	⋮	
j	ψ	
k	$\varphi \to \psi$	→I i–j

i	$\varphi \to \psi$	
j	φ	
k	ψ	→E i, j

The left schema is the rule →I (**implication-introduction**); the right one is →E (Definition 1.3.1). In →I, the line φ is an **assumption**, which opens the box, and the assumption is **discharged** when the box closes. The line ψ is a line **of that box itself**, not of a box inside it.

As in every schema (Definition 1.2.1), φ and ψ stand for any statements, replaced consistently, and i, j, k for line numbers. This rule has conditions that matter; read each one.

1. **Open a box** whose first line is the assumption φ. The assumption may be **any** statement: true or false, related to the premises or not. Its justification is "assumption", printed **"asm, for →I"**: the annotation names the rule that will close the box. The price of assuming anything we like is the box.
2. **Inside the box**, work as usual. Every rule may be used. A line inside may cite the assumption and every earlier line that is not inside a box that has already closed: the

premises, the lines above the box that are still in force, and the lines derived inside the box so far.

3. **The last line** of the box is ψ. It must be a line of this box itself, not a line of some further box opened inside it. (The drawing cannot show this, which is why the definition says it in words.)
4. **Close the box**, and write $\varphi \to \psi$ on a later line outside it, in practice the very next line. The justification cites the **whole box** as the range i–j, from its assumption to its last line, not its lines one by one.
5. **After the box closes**, nothing inside it may be cited again. The assumption is discharged, and everything derived under it goes with it. Only the implication leaves.

Clause 5 is the boundary of the temporary world. The subsection "Which lines a line may cite" below turns clause 5 into an exact rule.

What each box is for. From now on every assumption says, in its justification column, what its box is for. "asm, for →I" says that the box was opened to prove an implication, and that a →I line will close it. In a written proof the same line is the word "suppose". While a derivation is still growing and its closing line is not yet written, the annotation records the plan.

↻ **Second reading** · *A box of one line*

The **degenerate case** is $i = j$: a box of one line, whose assumption is also its last line. Then ψ is φ, and the rule writes $\varphi \to \varphi$. Nothing in the rule asks the box to contain more than its assumption.

Why the rule is right. The implication $\varphi \to \psi$ claims only that ψ holds whenever φ does. The box shows exactly that: starting from φ, and from facts that hold anyway, the rules lead to ψ.

Why the box closes. The lines inside hold only under the assumption. If the assumption is false, as $x > 2$ is at $x = 0$, they may be false too. Closing the box is what stops them being used where the assumption is not in force.

1.4.3 A first box: implication is transitive

The first result proved with a box is the chaining of implications. Section 1.3 used chains of implications forwards, firing each on a hypothesis already in hand. Here the hypothesis is not in hand, and we still want the chain.

Proposition 1.4.1. Implication Is Transitive

$P \to Q,\ Q \to R \vdash P \to R.$

Idea. Start from the goal. Its main connective is →, so the last line will be →I, and that needs a box that opens with P and ends with R. Inside the box, the new goal is R. Reason backwards: R is the conclusion of the premise $Q \to R$, so →E gives it once Q is on a line. And Q is the conclusion of $P \to Q$, so →E gives it from P, which the box has just assumed.

In the Givens/Goal table, →I moves the hypothesis into the Givens and makes the conclusion

the goal.

Givens	Goal
$P \to Q$, $Q \to R$	$P \to R$
$P \to Q$, $Q \to R$, P (box opened)	R
$P \to Q$, $Q \to R$, P, Q ($\to$E)	R
$P \to Q$, $Q \to R$, P, Q, R ($\to$E)	R ✓

The same search as a growing derivation: the box opened, its goal R at the bottom (left); then line 4 written forwards (right).

1	$P \to Q$	
2	$Q \to R$	
3	$\quad P$	asm, for $\to$I
	$\quad \vdots$	
5	$\quad R$	
6	$P \to R$	$\to$I 3–5

1	$P \to Q$	
2	$Q \to R$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q$	$\to$E 1, 3
	$\quad \vdots$	
5	$\quad R$	
6	$P \to R$	$\to$I 3–5

Proof. The following derivation has premises $P \to Q$ and $Q \to R$, and last line $P \to R$.

1	$P \to Q$	
2	$Q \to R$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q$	$\to$E 1, 3
5	$\quad R$	$\to$E 2, 4
6	$P \to R$	$\to$I 3–5

Line 3 opens a box with the assumption P. Lines 4 and 5 apply $\to$E inside it, each citing a premise and a line of the box. Line 5, R, is the last line of the box. Line 6 closes it and writes $P \to R$, citing the box 3–5. By Definition 1.2.3, this proves the proposition.

■

Read the drawing once more. The bar that starts at line 3 stops after line 5. The statements P, Q, and R on lines 3–5 were never established outright; they hold inside the temporary world where P is assumed. Only $P \to R$ stands outside, next to the premises.

Now give the letters some content. Let x be a real number, and let

$$P : x > 3, \qquad Q : x^2 > 9, \qquad R : x^2 > 4.$$

The premise $P \to Q$ holds: when $x > 3$, we have $x > 0$, so $x^2 = x \cdot x > 3x > 9$. The premise

$Q \to R$ holds because $4 < 9$: if $x^2 > 9$, then $x^2 > 4$. With the premises in hand, lines 3–6 become a paragraph, one sentence per line.

Line	Sentence
3	Suppose $x > 3$.
4	Since $x > 3$ implies $x^2 > 9$, we get $x^2 > 9$.
5	Since $x^2 > 9$ implies $x^2 > 4$, we get $x^2 > 4$.
6	Hence, if $x > 3$, then $x^2 > 4$.

A practiced writer would compress this to "if $x > 3$, then $x^2 > 9 > 4$": the same derivation with its joints hidden. When a proof is hard to follow, expanding it back into this form is how one finds the gap.

Remark. The same result holds with the two premises joined into one, $(P \to Q) \wedge (Q \to R) \vdash P \to R$. Two $\wedge$E lines take the premise apart first (premise-directed), and the rest is the derivation above, renumbered.

1	$(P \to Q) \wedge (Q \to R)$	
2	$P \to Q$	$\wedge$E 1
3	$Q \to R$	$\wedge$E 1
4	$\quad P$	asm, for $\to$I
5	$\quad Q$	$\to$E 2, 4
6	$\quad R$	$\to$E 3, 5
7	$P \to R$	$\to$I 4–6

1.4.4 Argument B, diagnosed

Section 1.1 broke the form of Argument B's last step. Here is the rule B misused. With E: "n is even" and S: "n^2 is even", B's prose is this derivation.

1	$\quad E$	asm, for $\to$I
2	$\quad S$	algebra 1
3	$E \to S$	$\to$I 1–2

Line 1 is B's "Suppose n is even, say $n = 2k$". Line 2 is its computation $n^2 = 4k^2 = 2(2k^2)$, a step of algebra, as in Example (A product of numbers in the unit interval). Line 3 is what the box earns; set it beside what B claimed.

what B's box earns		*what B claimed*
E / S		
$E \to S$	$\neq$	$S \to E$

B's box opens with E and closes on S, so $\to$I lets it write $E \to S$: the converse of the claim. B's last sentence wrote $S \to E$. The rule asks the box to open with the **hypothesis** of the implication being proved, here S, and B never opened that box.

The repair fails. Keep $E \to S$, open a box with S, and inside it use S to reach E: that step goes from $E \to S$ and S to E. It is affirming the consequent, the bad form of Section 1.3. Try to break it. Take "if $x > 2$, then $x^2 > 4$" and "$x^2 > 4$". At $x = -3$ both are true, since $9 > 4$, and "$x > 2$" is false. So B cannot be saved by moving its lines around: a $\to$I proof of $S \to E$ opens its box with S.

Warning. The box of $\to$I opens with the **hypothesis** of the implication to be proved, never with its conclusion: Argument B used $\to$I with the box opened at the wrong end.

Quick check. A proof of "if $x^3 > 8$, then $x > 2$" begins "Suppose $x > 2$." and ends "Hence $x^3 > 8$." What does its box earn? Is the claim it was meant to prove true?

Solution. Its box opens with $x > 2$ and closes on $x^3 > 8$, so it earns $(x > 2) \to (x^3 > 8)$, the converse of the claim. The claim itself is true: if $x \leq 2$, then $x^3 \leq 2^3 = 8$, since a larger real number has a larger cube, so $x^3 > 8$ forces $x > 2$. But this box does not prove it.

↻ **Second reading** · *Two errors that cancel*

Example. Two errors that cancel

Let x be a real number. Here is a "proof" that if $x > 3$, then $x^2 > 9$: *Suppose* $x^2 > 9$. *Then* $x > 3$. *Hence, if* $x > 3$, *then* $x^2 > 9$. Is the conclusion true? Is it proved?

Solution. As a derivation, the "proof" reads as follows.

1	$\quad x^2 > 9$	asm, for $\to$I
2	$\quad x > 3$	algebra 1
3	$(x > 3) \to (x^2 > 9)$	$\to$I 1–2

Two lines fail. Line 2 is no fact about numbers: at $x = -4$ we have $x^2 = 16 > 9$, and $x > 3$ fails. Line 3 writes the converse of what the box earns. The box opens with $x^2 > 9$ and ends with $x > 3$, so $\to$I would write $(x^2 > 9) \to (x > 3)$.

The two errors cancel. The implication the box would earn is false, and reversed it is the true claim: if $x > 3$, then $x > 0$, so $x^2 = x \cdot x > 3x > 9$. A reader who compares only the last line with the claim sees nothing wrong; a checker who reads every line finds both errors. A correct conclusion is no evidence that the lines above it are correct.

1.4.5 A box must end on a line of its own

Try to derive $\vdash P \to (Q \to P)$. The goal is an implication, so open a box with P; the new goal $Q \to P$ is again an implication, so open a box with Q inside the first. The statement P is already on line 1, so it is tempting to close the inner box at once.

1	P	asm, for →I
2	$\quad Q$	asm, for →I
3	$Q \to P$	→I 2–2
4	$P \to (Q \to P)$	→I 1–3

A checker reads the box 2–2: its assumption is Q, and its last line is Q. So →I may write only $Q \to Q$ from it, and line 3 is rejected. The statement P stands on line 1, outside the inner box, and clause 3 asks for ψ as a line **of the box itself**.

The fix is bookkeeping: copy P into the inner box. Copying an available line is the rule R (**repetition**), made official as clause (D3) of the definition below. The copy "P, R 1" then ends the inner box.

Some books let a box end on any line available inside it. This book copies the line in, so that every →I line can be checked by looking only at the box it cites. The corrected derivation is Example (First boxes) (b) below.

1.4.6 Which lines a line may cite

In a derivation without boxes (Definition 1.2.2), a line may cite any **earlier** line. With boxes that is no longer enough: after the derivation of Proposition 1.4.1, a line 7 citing line 5 would write R outside the box, the "$x^2 > 4$ for every x" error of the first subsection. What matters is which earlier lines are still **available**, and the bars show it.

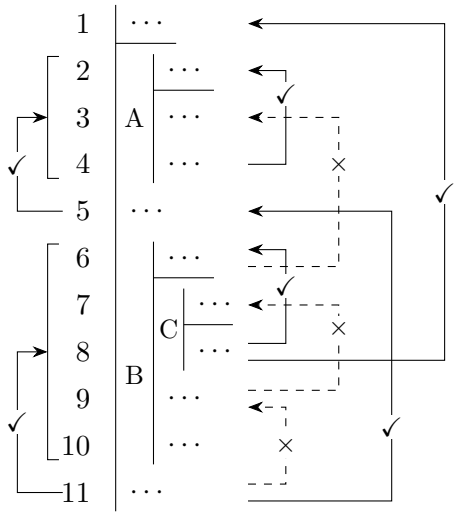
A line inside a box may be cited only from inside that box. Once a box has closed above you, nothing inside it can be cited; the boxes you are in may be left across their tops, and passing beside a closed box is fine.

Open a box, close a box: an assumption lives exactly as long as its box.

A **whole box**, cited by →I, has its own condition: the citing line must stand in exactly the boxes that the cited box stands in, other than the cited box itself. For single lines:

- a line in no box is available at every later line;
- a line inside a box is available further down that box, and inside every box opened within it;
- once its box closes, the line is available nowhere.

The figure shows eleven lines without statements: box A is lines 2–4, box B lines 6–10, and box C, lines 7–8, sits inside B. An arrow runs from a citing line up to the line it cites, solid and marked ✓ if legal, dashed and marked × if not. A bracket on the left marks a whole box cited.



A line can be cited from below only while every box containing it is still open. Line 8 leaves box C across its top to cite line 6, and line 11 passes beside the closed box B to cite line 5. Lines 5 and 11 stand in no box, and neither do the boxes A and B they cite whole.

Quick check. In the figure, which of lines 1–9 may line 10 cite?

Solution. Lines 1, 5, 6 and 9. Lines 1 and 5 are in no box; lines 6 and 9 are in box B, which contains line 10. Lines 2–4 are not available (box A has closed), and neither are lines 7–8 (box C has closed). It may also cite the whole box 7–8, which, like line 10, stands in box B alone.

Here is the official version, the one a checker uses.

Definition 1.4.2. Derivation

Fix a list of rules. A **derivation** is a finite list of numbered lines $1, 2, \ldots, n$, each carrying one statement and a justification, together with a set of **boxes**. A box is a block of consecutive lines whose first line is an assumption. Any two boxes are either disjoint or **one inside the other**. A box ends at one of its lines and is **closed** after it. The **last line** of the derivation lies outside every box.

The scope rule. Line i is **available** at line k if $i < k$ and **every** box containing line i also contains line k. A box is **available** at line k if it ends before line k and the boxes other than itself that contain it are **exactly** the boxes containing line k. A line may cite only lines and boxes available at it.

Every line is exactly one of the following:

(D1) a **premise**; the premises are the first lines, above every box, with a short rule under the last of them;

(D2) an **assumption**, which may be **any** statement, and opens a new box;

(D3) a **repetition** of the statement on a line available at this one, justified "R" and the line repeated;

(D4) the conclusion of **one** rule of the list, applied to lines and boxes available at this one,

such that this line and the cited lines and boxes form an instance of the rule's schema.

Clause (D3) is the rule R (**repetition**), whose schema is

$$\begin{array}{lll} i & \varphi & \\ & \vdots & \\ k & \varphi & \text{R } i \end{array}$$

where line i must be available at line k.

(D1)–(D4) say: given, assumed, copied, or one step of one rule, always from what is still available.

The scope rule is the definition's real content: line i is available at line k when every assumption that line i was derived under is still in force at line k: the picture rule above, in words a checker can apply.

Second reading · *Why checking stays mechanical*

Well defined. Whether a line is available is decided by looking at the boxes; nothing about the meaning of the statements enters. Checking a derivation stays mechanical.

Nothing earlier changes. A derivation with no boxes is a box-free derivation of Definition 1.2.2 for the same rules with R added: with no boxes, "available" means just "earlier". So every derivation of Sections 1.2 and 1.3 is a derivation in the new sense, and Definition 1.2.3 is read from now on with this definition. The sequent $\Gamma \vdash \psi$ still means: **there is** a derivation whose premises all belong to Γ and whose last line is ψ. Its last line lies in no box, so it depends on the premises and on nothing that was merely assumed.

Second reading · *More on repetition*

The typical use of R copies a line into a box opened below it.

$$\begin{array}{lll} i & \varphi & \\ & \vdots & \\ j & \quad \psi & \text{asm, for } \to\text{I} \\ & \quad \vdots & \\ k & \quad \varphi & \text{R } i \end{array}$$

Line i is available at line k exactly when every box containing line i contains line k; in this drawing, line i is in no box that has closed, so the copy is legal. The worked example Example (First boxes) below uses it.

Warning. A box that is never closed cannot end a derivation. If the last line of a list is still inside a box, it depends on an assumption that was never discharged, and Definition 1.4.2 does not count the list as a derivation at all. "Suppose $x > 2$. Then $x^2 > 4$", with nothing after it, proves nothing about x: the implication is the only thing the supposition was allowed to produce, and it was never written.

1.4.7 How to check a derivation

Section 1.2 checked a line by asking which rule it uses, which lines it cites, and whether they are earlier. With boxes the third question becomes: **are they available here?** For each line after the premises, ask:

1. **Which rule?** Name it. An assumption opens a box; R repeats; every other line applies one rule.
2. **Which lines or boxes does it cite, and do they form an instance of the rule's schema?** $\to$E cites an implication and exactly its hypothesis; $\to$I cites one box, whose assumption and last line must be the two sides of the implication written.
3. **Are they available here?** Apply the scope rule to each cited line and box.

Finding a derivation may need an idea; checking one never does. Here is the check run on the derivation of Proposition 1.4.1.

Line	Which rule?	Cites	Available?
1, 2	premises (D1)	—	—
3	assumption (D2); opens the box 3–5	—	—
4	$\to$E: Q from $P \to Q$ and P	1, 3	1 is in no box; 3 is in box 3–5, which contains 4
5	$\to$E: R from $Q \to R$ and Q	2, 4	2 is in no box; 4 is in box 3–5, which contains 5
6	$\to$I: $P \to R$ from the box	box 3–5	the box ended at 5; neither it nor 6 lies in another box

Every line passes, and the last line lies in no box. The bar shows the temporary world: it starts at line 3 and stops after line 5.

1.4.8 An assumption used outside its box

Now watch the scope rule catch an error. Take the derivation of Proposition 1.4.1 and add one careless line.

Example. A line cited from a closed box

Is the following a derivation? If it were, what would it show?

1	$P \to Q$	
2	$Q \to R$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q$	$\to$E 1, 3
5	$\quad R$	$\to$E 2, 4
6	$P \to R$	$\to$I 3–5
7	R	R 5

Solution. No. Lines 1–6 are the derivation of Proposition 1.4.1, and line 7 claims to be a repetition (D3). Line 5 lies inside the box 3–5, and that box does not contain line 7. By the scope rule, line 5 is not available at line 7, and (D3) repeats only an available line. So line 7 is illegal.

If it were allowed, the list would show $P \to Q,\ Q \to R \vdash R$, and that would be a disaster. Let x be a real number, and take P: "$x > 3$", Q: "$x > 2$", R: "$x > 1$". Both premises are true whatever x is. At $x = 0$ the conclusion $x > 1$ is false. A step that leads from true premises to a false conclusion cannot be a step of proof.

Warning. The assumption of a closed box is gone, and so is everything derived from it. Here is a typical slip. *"Suppose $x > 1$. Then $x^2 = x \cdot x > 1 \cdot x > 1$. Hence $x > 1$ implies $x^2 > 1$. Moreover $x^3 = x \cdot x^2 > 1$, so $x^3 > 1$."* The sentence "Hence ..." closed the box. The last sentence still uses $x > 1$ and $x^2 > 1$, and so it asserts $x^3 > 1$ with no hypothesis, which fails at $x = 0$. The repair is to keep the box open until everything that needs $x > 1$ is done, and to conclude "if $x > 1$, then $x^2 > 1$ and $x^3 > 1$".

Quick check. In the derivation below, which of lines 3 to 7 may cite line 2 individually?

1	$P \to Q$	
2	$\quad P$	asm, for $\to$I
3	$\quad Q$	$\to$E 1, 2
4	$\quad P \wedge Q$	$\wedge$I 2, 3
5	$\quad Q \wedge P$	$\wedge$I 3, 2
6	$P \to (Q \wedge P)$	$\to$I 2–5
7	$(P \to Q) \wedge (P \to (Q \wedge P))$	$\wedge$I 1, 6

Solution. Lines 3, 4, and 5. They lie in the box 2–5, the only box containing line 2, so line 2 is available to each of them. Line 6 cites the box 2–5 as a whole, which is what $\to$I requires; it does not cite line 2 by itself, and could not, since line 6 is outside the box. Line 7 may not

cite line 2 either, for the same reason.

1.4.9 More boxes: from nothing, and one inside another

Example. First boxes

Find a derivation of each sequent.

(a) $\vdash P \to P$.

(b) $\vdash P \to (Q \to P)$.

(c) $P \to (Q \to R) \vdash Q \to (P \to R)$.

Solution.

(a) *Idea.* The goal is an implication, so open a box with its hypothesis P; the box must end with P. The assumption is already P, so the box can end at once: a box of one line, whose assumption is also its last line. Nothing in →I asks a box to contain more than its assumption.

1	P	asm, for →I
2	$P \to P$	→I 1–1

Line 1 is an assumption (D2) and opens the box 1–1, whose last line is P. Line 2 cites that box, which ended before line 2 and, like line 2, lies in no other box. There are no premises, so this shows $\vdash P \to P$. It is the first statement in the book derivable from **no** premises by the rules of logic alone. Argument B's box in "Argument B, diagnosed" also began without premises, but it used a step of algebra. Every other rule so far (Definition 1.2.5, Definition 1.3.1, and R of Definition 1.4.2) cites at least one earlier line, so without a box a derivation must start with a premise. A box is what escapes: its first line is an assumption, and the assumption is discharged before the end.

(b) *Idea.* This is the derivation the checker accepted after the fix in "A box must end on a line of its own": line 3 copies P into the inner box.

1	P	asm, for →I
2	Q	asm, for →I
3	P	R 1
4	$Q \to P$	→I 2–3
5	$P \to (Q \to P)$	→I 1–4

Line 3 is legal: the only box containing line 1, the box 1–4, also contains line 3. The boxes are nested, and the inner one closes first. No line cites line 2 on its own: the assumption Q is never used, and nothing in →I requires it to be. An implication whose conclusion is already known holds whatever its hypothesis.

(c) *Idea.* The goal $Q \to (P \to R)$ asks for a box with Q and, inside it, a box with P ending in R. Inside both boxes, work from the premise: its main connective is the first →, with hypothesis P, so →E fires on the assumption P and gives $Q \to R$. That

implication fires on the assumption Q and gives R, the inner goal.

Both boxes opened, goals at their feet (left); then the inner lines written, and only line 7 still to justify (right).

1	$P \to (Q \to R)$	
2	$\quad Q$	asm, for $\to$I
3	$\quad\quad P$	asm, for $\to$I
	$\quad\quad \vdots$	
5	$\quad\quad R$	
6	$\quad P \to R$	
7	$Q \to (P \to R)$	

1	$P \to (Q \to R)$	
2	$\quad Q$	asm, for $\to$I
3	$\quad\quad P$	asm, for $\to$I
4	$\quad\quad Q \to R$	$\to$E 1, 3
5	$\quad\quad R$	$\to$E 4, 2
6	$\quad P \to R$	$\to$I 3–5
7	$Q \to (P \to R)$	?

The finished derivation:

1	$P \to (Q \to R)$	
2	$\quad Q$	asm, for $\to$I
3	$\quad\quad P$	asm, for $\to$I
4	$\quad\quad Q \to R$	$\to$E 1, 3
5	$\quad\quad R$	$\to$E 4, 2
6	$\quad P \to R$	$\to$I 3–5
7	$Q \to (P \to R)$	$\to$I 2–6

Line 5 cites line 2 from inside the inner box: every box containing line 2 (only 2–6) contains line 5. So the order of two hypotheses does not matter. For real numbers a and x, the derivation turns "if $a > 0$, then (if $x > 1$, then $ax > a$)" into "if $x > 1$, then (if $a > 0$, then $ax > a$)".

In (b) and (c) the goal told us which boxes to open, and in which order: the outermost $\to$ of the goal gives the outermost box. This is the goal-directed move of this section. *An implication goal opens a box with its hypothesis; the new goal is its conclusion.*

1.4.10 Two hypotheses, or one at a time

A theorem with two hypotheses, "if $x > 0$ and $y > 0$, then $xy > 0$", is often written with one hypothesis at a time: "let $x > 0$; then, if $y > 0$, $xy > 0$". The two forms are $(P \wedge Q) \to R$ and $P \to (Q \to R)$. Each can be derived from the other, so a writer may use whichever reads better.

Proposition 1.4.2. Two Hypotheses or One at a Time

$(P \wedge Q) \to R \vdash P \to (Q \to R)$, and $P \to (Q \to R) \vdash (P \wedge Q) \to R$.

Try it first (ten minutes): for the first claim, which boxes will you open, and in which order?

Idea. The tempting first move is premise-directed: use $(P \wedge Q) \to R$ at once. But $\to$E needs $P \wedge Q$ on a line, and at the start nothing is on a line except the premise. So start from the goal. Its main connective is the first $\to$, so open a box with P; the new goal $Q \to R$ is again an implication, so open a box with Q inside the first. Now both P and Q are available, $\wedge$I builds $P \wedge Q$, and the premise fires. For the second claim the goal $(P \wedge Q) \to R$ asks for one box, with $P \wedge Q$; take the assumption apart and feed the halves to the premise one at a time, P first because it is the premise's outer hypothesis.

The first claim: the outer box opened (left), then the inner one, with goal R (right).

1	$(P \wedge Q) \to R$	
2	$\quad P$	asm, for $\to$I
	$\quad \vdots$	
6	$\quad Q \to R$	
7	$P \to (Q \to R)$	$\to$I 2–6

1	$(P \wedge Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad Q$	asm, for $\to$I
	$\quad\quad \vdots$	
5	$\quad\quad R$	
6	$\quad Q \to R$	$\to$I 3–5
7	$P \to (Q \to R)$	$\to$I 2–6

The second claim: the box opened (left), then the assumption taken apart (right).

1	$P \to (Q \to R)$	
2	$\quad P \wedge Q$	asm, for $\to$I
	$\quad \vdots$	
6	$\quad R$	
7	$(P \wedge Q) \to R$	$\to$I 2–6

1	$P \to (Q \to R)$	
2	$\quad P \wedge Q$	asm, for $\to$I
3	$\quad P$	$\wedge$E 2
4	$\quad Q$	$\wedge$E 2
	$\quad \vdots$	
6	$\quad R$	
7	$(P \wedge Q) \to R$	$\to$I 2–6

Before reading on: in the first derivation, which lines may line 4 cite, and why may line 6 cite the box 3–5 but not line 5 on its own?

Proof. For the first claim, the following is a derivation with premise $(P \wedge Q) \to R$ and last line $P \to (Q \to R)$.

1	$(P \wedge Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad Q$	asm, for $\to$I
4	$\quad\quad P \wedge Q$	$\wedge$I 2, 3
5	$\quad\quad R$	$\to$E 1, 4
6	$\quad Q \to R$	$\to$I 3–5
7	$P \to (Q \to R)$	$\to$I 2–6

Line 4 may cite lines 1, 2, and 3: line 1 lies in no box, the only box containing line 2 (the box 2–6) contains line 4, and both boxes containing line 3 contain line 4. Line 6 may cite the box 3–5, because that box has ended and both it and line 6 lie in the box 2–6 and in no other. Line 6 may not cite line 5 on its own, because the box 3–5 contains line 5 and not line 6.

For the second claim, the following is a derivation with premise $P \to (Q \to R)$ and last line $(P \wedge Q) \to R$.

1	$P \to (Q \to R)$	
2	$\quad P \wedge Q$	asm, for $\to$I
3	$\quad P$	$\wedge$E 2
4	$\quad Q$	$\wedge$E 2
5	$\quad Q \to R$	$\to$E 1, 3
6	$\quad R$	$\to$E 5, 4
7	$(P \wedge Q) \to R$	$\to$I 2–6

Lines 3–6 cite line 1, which lies in no box, or earlier lines of the box 2–6, from inside that box. Line 7 cites the box 2–6, which has ended, and neither that box nor line 7 lies in any other box. By Definition 1.2.3, the two derivations prove the two claims.

■

Quick check. In the second derivation of Proposition 1.4.2, what does the annotation of line 2 say, and which line closes its box?

Solution. "asm, for $\to$I": the box was opened to prove an implication. Line 7 closes it, citing the box 2–6.

In the first derivation the boxes are nested, and the inner one closes first: an assumption made later is discharged earlier. That is the general pattern, forced by Definition 1.4.2, since two boxes are either disjoint or one inside the other.

Read mathematically, the first claim turns "if $x > 0$ and $y > 0$, then $xy > 0$" into "if $x > 0$,

then (if $y > 0$, then $xy > 0$)", and the second turns it back. The two forms say the same thing, one hypothesis at a time or both at once.

Quick check. To show $Q \wedge R \vdash (P \to Q) \wedge (S \to R)$, does the derivation need one box or two?

Solution. Two. The goal's main connective is $\wedge$, which opens no box, so the goal splits into the two implications $P \to Q$ and $S \to R$ (goal-directed $\wedge$I). Each is proved by $\to$I, with its own box, opened by its own hypothesis. Inside each box, $\wedge$E on the premise gives the half that box needs.

1	$Q \wedge R$	
2	$\quad P$	asm, for $\to$I
3	$\quad Q$	$\wedge$E 1
4	$P \to Q$	$\to$I 2–3
5	$\quad S$	asm, for $\to$I
6	$\quad R$	$\wedge$E 1
7	$S \to R$	$\to$I 5–6
8	$(P \to Q) \wedge (S \to R)$	$\wedge$I 4, 7

The two boxes are disjoint: the first closes before the second opens. Line 8 may cite lines 4 and 7, which lie in no box.

1.4.11 The rule card

Before the card, the rule at work on large statements. Each row is one use of $\to$I: the box, and the line it writes.

Example. Instances of $\to$I

φ	ψ	The box	New line
$P \wedge \neg Q$	$P \wedge (P \wedge \neg Q)$	from $P \wedge \neg Q$ (asm) to $P \wedge (P \wedge \neg Q)$, by $\wedge$E and then $\wedge$I	$(P \wedge \neg Q) \to (P \wedge (P \wedge \neg Q))$
$Q \to R$	$(P \to Q) \to (P \to R)$	from $Q \to R$ (asm) to $(P \to Q) \to (P \to R)$, with two boxes inside it	$(Q \to R) \to ((P \to Q) \to (P \to R))$
$(P \vee Q) \wedge R$	$(P \vee Q) \wedge R$	a box of one line	$((P \vee Q) \wedge R) \to ((P \vee Q) \wedge R)$

Not an instance. A box opens with Q and ends with P, copied in by R. The line "$P \to Q$ — $\to$I" is not an instance: $\to$I writes the assumption first, $Q \to P$. This is Argument B's step.

The card grows by two rows. In *Shape*, "box $\varphi \ldots \psi$" is a closed box from the assumption φ to its last line ψ.

Rule	Shape	You write	Lean
$\wedge$I	$\varphi, \psi \Longrightarrow \varphi \wedge \psi$	"We show P, and we show Q."	`constructor`
$\wedge$E	$\varphi \wedge \psi \Longrightarrow \varphi$ (or ψ)	"In particular, P."	`h.1` (and `h.2`)
$\to$E	$\varphi \to \psi, \varphi \Longrightarrow \psi$	"Since P implies Q, and P holds, Q."	`exact h hp`
$\to$I	box $\varphi \ldots \psi \Longrightarrow \varphi \to \psi$	"Assume P. … Hence Q."	`intro hp`
R	φ (available) $\Longrightarrow \varphi$, copied	"As shown above, P."	`exact hp`

In the Lean column, `intro hp` opens the box and names its assumption `hp`.

Warning (Not allowed: $\to$I). Each step below is wrong; the tag names the clause of the rule's reading that it breaks.

- Close the box on a line of a box inside it: clause 3, the last line must be a line of the box itself ("A box must end on a line of its own").
- Write $\psi \to \varphi$ from a box that runs from φ to ψ: clause 4, the box opens with the hypothesis (Argument B).
- Cite a line inside a closed box: clause 5 (Example (A line cited from a closed box)).
- End the derivation inside an open box: clause 4, the box must close (the warning on a box that is never closed).
- Cite the box's lines one by one: clause 4, the box is cited whole, as a range.

The discharge is where the rule bites, so here is one wrong step beside the right one, with premises $(P \wedge R) \to Q$ and R. On the left, line 6 carries Q out of the box by repetition, as if P had been proved. On the right, line 6 closes the box with $\to$I.

1	$(P \wedge R) \to Q$	
2	R	
3	$\quad P$	asm, for $\to$I
4	$\quad P \wedge R$	$\wedge$I 3, 2
5	$\quad Q$	$\to$E 1, 4
6	Q	R 5

1	$(P \wedge R) \to Q$	
2	R	
3	$\quad P$	asm, for $\to$I
4	$\quad P \wedge R$	$\wedge$I 3, 2
5	$\quad Q$	$\to$E 1, 4
6	$P \to Q$	$\to$I 3–5

The left list would show $(P \wedge R) \to Q,\ R \vdash Q$. Let x be a real number, with P: "$x > 5$", R: "$x > 0$", Q: "$x > 3$". The first premise holds for every x, and at $x = 1$ both premises are true and Q is false.

The sentence for $\to$I is the one to watch. "Assume P" opens a box, and "hence Q" is its last line. The closing line, "so P implies Q", is usually left unsaid in written mathematics. It is still there: the box closes whether or not anyone says so. It is also the line Argument B got wrong.

Remark (Looking ahead). Chapter 2 proves the claim Argument B wanted, with the rules for "not".

1.4.12 Exercises

1.4.12.1 A. Check your understanding

Exercise. A1

Consider the following derivation, whose justifications are missing.

1	$P \to (Q \to R)$	
2	Q	
3	$\quad P$	asm, for $\to$I
4	$\quad Q \to R$	(i)
5	$\quad R$	(ii)
6	$P \to R$	(iii)

(a) Write the schema of $\to$I from memory, and state the condition that the drawing cannot show.

(b) Fill in the justifications (i), (ii), (iii), and say what the derivation shows, written with $\vdash$.

(c) Which of lines 1–5 are available at line 6, and which boxes?

(d) Someone extends the derivation by a line 7, the statement R justified by $\to$E 6, 3. Is the result a derivation? If not, name the clause of Definition 1.4.2 that line 7 breaks.

(e) In the first derivation in the proof of Proposition 1.4.2, why is line 4 available at line 5? And why is the inner box closed before the outer one?

(f) In the same derivation, which rule produces line 5, and which lines does it cite?

Solution.

(a) The schema is the left one in Definition 1.4.1: a box from an assumption φ on line i to a line ψ on line j, and then $\varphi \to \psi$ on line k, justified $\to$I i–j. The condition the drawing cannot show: ψ is a line of that box itself, not of a box inside it.

(b) (i) $\to$E 1, 3; (ii) $\to$E 4, 2; (iii) $\to$I 3–5. The derivation shows $P \to (Q \to R),\ Q \vdash P \to R$.

(c) Lines 1 and 2, and the box 3–5. Lines 3, 4, and 5 lie in the box 3–5, which does not contain line 6, so none of them is available there one by one. The box as a whole is available at line 6: it has ended, and neither it nor line 6 lies in another box. That is why (iii) may cite it.

(d) No. Line 6 is available at line 7, but line 3 is not: it lies in the box 3–5, which does not contain line 7. So line 7 applies a rule to a line that is not available, which (D4) forbids.

(e) Line 4 lies in two boxes, 2–6 and 3–5, and both contain line 5; so by the scope rule line

4 is available at line 5. The inner box 3–5 lies inside the outer box 2–6, and two boxes are either disjoint or one inside the other, so the inner box must end first. Read off the goal: the inner box proves $Q \to R$, which is the last line the outer box needs.

(f) →E, citing line 1, the implication $(P \wedge Q) \to R$, and line 4, its hypothesis $P \wedge Q$.

Exercise. A2: Drill

The parts run from one step, through a missing line, a justification column and steps to judge, to short derivations. Parts (e)–(g) refer to the eleven-line figure in "Which lines a line may cite".

(a) A box runs from line 3, $(P \vee Q) \wedge R$, to line 6, $S \to P$, a line of that box itself. What does "→I 3–6" write?

(b) Line 7 is $(P \to Q) \to R$, justified "→I 2–6". What is line 2, and what does its annotation read?

(c) Line 9 is $(P \wedge Q) \to (R \vee S)$, justified "→I 4–8". What are lines 4 and 8?

(d) Say which lines below are premises and which are assumptions, and fill in (i) and (ii).

1	P	
2	$\quad Q$	asm, for →I
3	$\quad P \wedge Q$	(i)
4	$Q \to (P \wedge Q)$	(ii)

(e) May line 8 cite line 5?

(f) May line 9 cite the whole box 2–4?

(g) May line 11 cite line 6?

(h) A box runs from line 3, P (an assumption), to line 6, R; line 4, inside it, is Q. Is "7. $P \to Q$ — →I 3–4" correct?

(i) Find the illegal line, and say why it is illegal.

1	$P \to Q$	
2	$\quad P$	asm, for →I
3	$\quad Q$	→E 1, 2
4	$P \to Q$	→I 2–3
5	Q	→E 4, 2

(j) Let x be a real number. Is "Suppose $x > 3$. Then $x^2 > 9$.", with nothing after it, a proof that $x^2 > 9$?

(k) Show $\vdash (P \wedge Q) \to (Q \wedge P)$ by a derivation. Then write the argument as a paragraph of prose, one sentence per line.

(l) Derive $P \to Q \vdash (P \wedge R) \to Q$.

(m) Derive $Q \vdash (P \wedge R) \to (Q \wedge R)$.

(n) Derive $\vdash (P \wedge Q) \to (R \to P)$. Which box opens first?

(o) Derive $(P \to Q) \wedge R \vdash P \to (Q \wedge R)$.

Solution.

(a) $((P \vee Q) \wedge R) \to (S \to P)$: the assumption on the left, the box's last line on the right.

(b) Line 2 is $P \to Q$, the hypothesis of line 7, and its annotation reads "asm, for $\to$I".

(c) Line 4 is the assumption $P \wedge Q$, which opens the box; line 8 is $R \vee S$, the last line of that box.

(d) Line 1 is a premise and line 2 an assumption; (i) $\wedge$I 1, 2; (ii) $\to$I 2–3.

(e) Yes: line 5 lies in no box.

(f) No: box 2–4 stands in no box, and line 9 stands in box B, so the two do not stand in exactly the same boxes.

(g) No: line 6 is inside box B, which closed after line 10.

(h) No: lines 3–4 are not a box, since the box runs on to line 6. $\to$I cites the whole box 3–6 and writes $P \to R$. ($P \to Q$ is earned by a box of its own that ends on Q.)

(i) Line 5: it cites line 2, which lies inside the box 2–3, closed before line 5.

(j) No. The box opened by "suppose" never closes, so nothing has left it; and at $x = 0$ the sentence $x^2 > 9$ is false. What the two sentences can prove, once the box closes, is "if $x > 3$, then $x^2 > 9$".

(k) The goal is an implication, so open a box with $P \wedge Q$; inside, $\wedge$E gives both halves, and $\wedge$I rebuilds them in the other order.

1	$P \wedge Q$	asm, for $\to$I
2	P	$\wedge$E 1
3	Q	$\wedge$E 1
4	$Q \wedge P$	$\wedge$I 3, 2
5	$(P \wedge Q) \to (Q \wedge P)$	$\to$I 1–4

In prose: *Assume P and Q. In particular, P holds. In particular, Q holds. Hence Q and P both hold. This proves that if P and Q, then Q and P.* Compared with Proposition 1.2.1, the premise $P \wedge Q$ has become an assumption, and the conclusion has become an implication that depends on nothing.

(l) 2 $P \wedge R$, assumption, for $\to$I; 3 P, $\wedge$E 2; 4 Q, $\to$E 1, 3; 5 $(P \wedge R) \to Q$, $\to$I 2–4.

(m) 2 $P \wedge R$, assumption, for $\to$I; 3 R, $\wedge$E 2; 4 $Q \wedge R$, $\wedge$I 1, 3; 5 $(P \wedge R) \to (Q \wedge R)$, $\to$I 2–4.

(n) The box of $P \wedge Q$ opens first, since the goal's main connective is its first $\to$: 1 $P \wedge Q$, assumption, for $\to$I; 2 R, assumption, for $\to$I, opening a box inside the first; 3 P, $\wedge$E 1; 4 $R \to P$, $\to$I 2–3; 5 $(P \wedge Q) \to (R \to P)$, $\to$I 1–4.

(o) 2 P, assumption, for $\to$I; 3 $P \to Q$, $\wedge$E 1; 4 Q, $\to$E 3, 2; 5 R, $\wedge$E 1; 6 $Q \wedge R$, $\wedge$I 4, 5; 7 $P \to (Q \wedge R)$, $\to$I 2–6.

1.4.12.2 B. Practice

Exercise. B1: Two boxes, one inside the other

A search for $(P \wedge Q) \to R,\ S \to Q \vdash S \to (P \to R)$ has reached the stage below. *Name the move, then complete it.*

1	$(P \wedge Q) \to R$	
2	$S \to Q$	
3	$\quad S$	asm, for $\to$I
4	$\quad\quad P$	asm, for $\to$I
	$\quad\quad \vdots$	
	$\quad \vdots$	
9	$S \to (P \to R)$	?

Solution. The move is goal-directed, twice: an implication goal opens a box with its hypothesis, and the new goal is its conclusion. The goal $S \to (P \to R)$ gave the outer box, with S, and its conclusion $P \to R$ gave the inner box, with P; the inner goal is R. Inside both boxes, work backwards from the first premise: R is its conclusion, so it suffices to build its hypothesis $P \wedge Q$ exactly. P is line 4, and Q comes from the second premise and the assumption S.

1	$(P \wedge Q) \to R$	
2	$S \to Q$	
3	$\quad S$	asm, for $\to$I
4	$\quad\quad P$	asm, for $\to$I
5	$\quad\quad Q$	$\to$E 2, 3
6	$\quad\quad P \wedge Q$	$\wedge$I 4, 5
7	$\quad\quad R$	$\to$E 1, 6
8	$\quad P \to R$	$\to$I 4–7
9	$S \to (P \to R)$	$\to$I 3–8

Every citation is available. Lines 1 and 2 lie in no box. Line 3 lies only in the box 3–8, which contains line 5. Lines 4, 5 and 6 lie in both boxes, which contain lines 6 and 7. Line 8 cites the box 4–7, which has ended and stands, like line 8, in the box 3–8 and no other. Line 9 cites the box 3–8, which has ended and, like line 9, lies in no other box.

Exercise. B2: An assumption that is not used

(a) Show that $Q \vdash P \to Q$, and say which line needs repetition.

(b) Show that $\vdash (P \wedge Q) \to (R \to Q)$.

(c) Let x be a real number. Write the instance of (a) with Q: "$x^2 \geq 0$" and P: "$x > 5$" as an English sentence, and say which hypothesis its proof never uses.

Solution.

(a) The goal is an implication, so open a box with P; its last line must be Q.

1	Q	
2	$\quad P$	asm, for $\to$I
3	$\quad Q$	R 1
4	$P \to Q$	$\to$I 2–3

Line 3 needs repetition: $\to$I wants Q as a line of the box 2–3 itself. It is legal because line 1 lies in no box.

(b) Two implications, so two nested boxes: one with $P \wedge Q$, and inside it one with R.

1	$P \wedge Q$	asm, for $\to$I
2	Q	$\wedge$E 1
3	$\quad R$	asm, for $\to$I
4	$\quad Q$	R 2
5	$R \to Q$	$\to$I 3–4
6	$(P \wedge Q) \to (R \to Q)$	$\to$I 1–5

Line 4 is legal: the only box containing line 2 also contains line 4. (Applying $\wedge$E to line 1 inside the inner box works as well.)

(c) "$x^2 \geq 0$; therefore, if $x > 5$, then $x^2 \geq 0$." The hypothesis $x > 5$ is never used: its box only repeats what was known.

Exercise. B3: Unwinding an absolute value

(a) Show that $P \to Q,\ P \to R \vdash P \to (Q \wedge R)$.

(b) Let x be a real number. Use (a) to deduce "if $|x-2| < 1$, then $1 < x < 3$" from the two facts "if $|x-2| < 1$, then $1 < x$" and "if $|x-2| < 1$, then $x < 3$". Write the resulting argument as prose.

Solution.

(a) The goal is an implication, so assume P, and reach $Q \wedge R$ by getting each half.

1	$P \to Q$	
2	$P \to R$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q$	$\to$E 1, 3
5	$\quad R$	$\to$E 2, 3
6	$\quad Q \wedge R$	$\wedge$I 4, 5
7	$P \to (Q \wedge R)$	$\to$I 3–6

The assumption on line 3 is cited twice, by lines 4 and 5: a line is not used up when it is cited.

(b) Take P: "$|x-2| < 1$", Q: "$1 < x$" and R: "$x < 3$". The two facts are the premises $P \to Q$ and $P \to R$, and $Q \wedge R$ is $(1 < x) \wedge (x < 3)$, which $1 < x < 3$ abbreviates

(Section 1.1). The derivation in (a), with these statements put in for the letters, is a derivation of $P \to (Q \wedge R)$ from the two facts.

In prose: *Suppose* $|x - 2| < 1$. *Then* $1 < x$ *by the first fact, and* $x < 3$ *by the second. Hence* $1 < x < 3$. *This proves that* $|x - 2| < 1$ *implies* $1 < x < 3$.

Exercise. B4: Putting a derivation in order

Here are nine lines, each a statement with its justification. Seven of them, in the right order and with the right boxes, form a derivation of $P \to (Q \to R) \vdash Q \to (P \to R)$. Give the order, and say where each box opens and closes. Then say why each of the other two lines cannot be used.

(i) $P \to R$, by $\to$I on the box from P to R.
(ii) Q, assumption, for $\to$I.
(iii) R, by $\to$E from $Q \to R$ and Q.
(iv) R, by R, repeating the line R of (iii), written directly below (i).
(v) $P \to (Q \to R)$, premise.
(vi) $Q \to (P \to R)$, by $\to$I on the box from Q to $P \to R$.
(vii) R, assumption, opening the box for the $\to$I that proves $P \to R$.
(viii) $Q \to R$, by $\to$E from $P \to (Q \to R)$ and P.
(ix) P, assumption, for $\to$I.

Solution. The order is (v), (ii), (ix), (viii), (iii), (i), (vi), with one box opened by (ii) and closed after (i), and a second box inside it, opened by (ix) and closed after (iii). This is the derivation of Example (First boxes) (c).

It is the only admissible order. The premise (v) comes first. The last line (vi) needs a box from Q to $P \to R$, so (ii) opens a box whose last line is (i). Line (i) needs a box from P to R, which must lie inside the box of Q, since (i) is a line of that box; so (ix) opens the inner box, whose last line is (iii). Inside it, (viii) needs the premise and (ix), and (iii) needs (viii) and (ii). Opening the box of P first and the box of Q inside it would not do: the inner $\to$I would then give $Q \to R$, and the outer one $P \to (Q \to R)$, the premise again.

(iv) cannot be used where it is placed. Directly below (i), the inner box has closed, and line (iii) lies inside it; so (iii) is not available there, and (D3) repeats only an available line (the scope rule of Definition 1.4.2). It is the error of Example (A line cited from a closed box).

(v) cannot be used. $\to$I proves $P \to R$ from a box opened by P, the **hypothesis**. A box opened by R would prove an implication with hypothesis R: the converse error of Argument B.

1.4.12.3 C. Going deeper

Exercise. C1: Exit check: a line from a closed box

A student wants to show $(P \wedge Q) \to R \vdash (P \to (Q \to R)) \wedge (Q \to (P \to R))$, and offers the

following derivation. Lines 1–7 are the first derivation in the proof of Proposition 1.4.2.

1	$(P \wedge Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad Q$	asm, for $\to$I
4	$\quad\quad P \wedge Q$	$\wedge$I 2, 3
5	$\quad\quad R$	$\to$E 1, 4
6	$\quad Q \to R$	$\to$I 3–5
7	$P \to (Q \to R)$	$\to$I 2–6
8	$\quad Q$	asm, for $\to$I
9	$\quad\quad P$	asm, for $\to$I
10	$\quad\quad R$	R 5
11	$\quad P \to R$	$\to$I 9–10
12	$Q \to (P \to R)$	$\to$I 8–11
13	$(P \to (Q \to R)) \wedge (Q \to (P \to R))$	$\wedge$I 7, 12

(a) Find the first illegal line, and name the clause of Definition 1.4.2 that it breaks.

(b) Is the student's conclusion derivable from the premise? Justify your answer.

(c) The student defends line 10: "at line 10, both P and Q are assumed, exactly as they were at line 5, so R holds there too." Explain why Definition 1.4.2 does not accept this defense, and show by an instance with real numbers that the move "repeat a line of a closed box" cannot be allowed.

Solution (Hint 1). For (a), check each line's citations against the scope rule; for (b), keep the student's idea and change one step.

Solution (Hint 2). For (b), inside the box of P at line 9, the assumptions Q (line 8) and P (line 9) are both available. What did lines 4 and 5 do with such a pair?

Solution.

(a) Line 10 is the first illegal line. It claims to repeat line 5 by (D3), but line 5 lies in the boxes 2–6 and 3–5, and neither contains line 10. So line 5 is not available at line 10, and (D3) allows repeating only an available line. Lines 1–9 are correct, and lines 11–13 would be correct if line 10 were.

(b) Yes: the sequent is derivable. Replace line 10 by two lines that do inside the new boxes what lines 4 and 5 did inside the old ones.

1	$(P \wedge Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad Q$	asm, for $\to$I
4	$\quad\quad P \wedge Q$	$\wedge$I 2, 3
5	$\quad\quad R$	$\to$E 1, 4
6	$\quad Q \to R$	$\to$I 3–5
7	$P \to (Q \to R)$	$\to$I 2–6
8	$\quad Q$	asm, for $\to$I
9	$\quad\quad P$	asm, for $\to$I
10	$\quad\quad P \wedge Q$	$\wedge$I 9, 8
11	$\quad\quad R$	$\to$E 1, 10
12	$\quad P \to R$	$\to$I 9–11
13	$Q \to (P \to R)$	$\to$I 8–12
14	$(P \to (Q \to R)) \wedge (Q \to (P \to R))$	$\wedge$I 7, 13

Every box containing line 8 or line 9 contains line 10, and line 1 lies in no box; the other lines check as before. The student's derivation was wrong and its conclusion right: a flawed argument can have a true conclusion.

(c) Availability is about **boxes**, not about which statements happen to be assumed. The boxes 8–11 and 9–10 of the student's derivation are new boxes; that their assumptions repeat those of the old boxes gives line 10 no access to what was derived inside the old ones. And the move, once allowed, can be used anywhere. Used outside every box, as a line 8 "R, R 5" after lines 1–7, it would derive $(P \wedge Q) \to R \vdash R$. Let x be a real number, and take P: "$x > 2$", Q: "$x > 3$" and R: "$x > 1$". The premise "if $x > 2$ and $x > 3$, then $x > 1$" is true whatever x is, and at $x = 0$ the conclusion $x > 1$ is false. A move that leads from a true premise to a false conclusion cannot be a step of proof.

Marking scheme. (10 points)

- 3: names line 10, because line 5 lies in boxes that do not contain it.
- 3: gives any correct derivation of the sequent, every citation available.
- 1: says the conclusion is derivable although the argument is wrong.
- 3: explains that availability depends on boxes, not on repeated assumptions, with an instance where the move leads from a true premise to a false conclusion.

Fatal errors (0 for the step): naming line 5 or line 9; claiming the sequent is not derivable; refuting the conclusion instead of the move.

Exercise. C2: A hypothesis inside a hypothesis

Show that $(P \to Q) \to (P \to R) \vdash P \to (Q \to R)$. Then say, in one sentence, why the premise cannot be used until a third box is opened.

Solution (Hint 1). Read the boxes off the goal. The premise's hypothesis is $P \to Q$: what would you need to derive it?

Solution (Hint 2). Start: line 2 assumes P and line 3 assumes Q, read off the goal. The last line of the box of Q must be R, and line 1 gives $P \to R$ only once $P \to Q$ is on an available line.

Solution. The goal $P \to (Q \to R)$ asks for a box with P and, inside it, a box with Q whose last line is R. The premise gives $P \to R$ as soon as we have $P \to Q$, and $P \to R$ with the available P gives R. So the task is to derive $P \to Q$ inside the box of Q, and that takes a third box.

1	$(P \to Q) \to (P \to R)$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad Q$	asm, for $\to$I
4	$\quad\quad\quad P$	asm, for $\to$I
5	$\quad\quad\quad Q$	R 3
6	$\quad\quad P \to Q$	$\to$I 4–5
7	$\quad\quad P \to R$	$\to$E 1, 6
8	$\quad\quad R$	$\to$E 7, 2
9	$\quad Q \to R$	$\to$I 3–8
10	$P \to (Q \to R)$	$\to$I 2–9

Line 5 is a legal repetition: both boxes containing line 3 contain line 5. Line 8 may cite line 2 because the box 2–9, the only box containing line 2, contains line 8.

The sentence: the premise fires only on an available line $P \to Q$, and no available line is $P \to Q$ or has it as a conjunct or as the conclusion of an implication, so it can come only from $\to$I, which needs a box of its own.

Line 1 does not help, since $\to$E on it would need $P \to Q$ already. The third box assumes P again although line 2 already did: $\to$I needs the box, not new information.

Marking scheme. (10 points)

- 2: opens a box with P and inside it a box with Q, read off the goal.
- 3: derives $P \to Q$ inside the box of Q by a box whose last line is Q.
- 2: obtains $P \to R$, then R, by $\to$E on available lines.
- 2: closes the boxes inner first, each $\to$I citing an available box.
- 1: the sentence says $P \to Q$ must be a line that no elimination rule can supply, hence needs a box.

Fatal errors (0 for the step): citing a line outside the box that contains it; writing $P \to Q$ with no box.

Exercise. C3: Exit check: a derivation and its prose

(a) Write a derivation of $P \to Q,\ (Q \to R) \wedge (R \to S) \vdash P \to S$, with every line justified.

(b) Let x be a real number, and take P: "$x > 5$", Q: "$x > 4$", R: "$x^2 > 16$", S: "$x^2 > 9$". Say why each premise is true for every real x. Then write your derivation as fluent prose about x, one sentence per step.

Solution.

(a) Premise-directed first: the second premise is a conjunction, so $\wedge$E takes it apart. Goal-directed next: the goal is an implication, so open a box with its hypothesis P, and reach S by firing the three implications in turn.

1	$P \to Q$	
2	$(Q \to R) \wedge (R \to S)$	
3	$Q \to R$	$\wedge$E 2
4	$R \to S$	$\wedge$E 2
5	$\quad P$	asm, for $\to$I
6	$\quad Q$	$\to$E 1, 5
7	$\quad R$	$\to$E 3, 6
8	$\quad S$	$\to$E 4, 7
9	$P \to S$	$\to$I 5–8

Lines 1–4 lie in no box, so they are available inside the box 5–8; line 9 cites that box, which has ended, and neither it nor line 9 lies in another box.

(b) The premise $P \to Q$ holds because $5 > 4$: if $x > 5$, then $x > 4$. For $Q \to R$: if $x > 4$, then $x > 0$, so $x^2 = x \cdot x > 4x > 16$. For $R \to S$: if $x^2 > 16$, then $x^2 > 9$, since $16 > 9$. So both premises (the second a conjunction of two true implications) hold for every real x.

In prose: *We know that $x > 5$ implies $x > 4$, and that $x > 4$ implies $x^2 > 16$ and $x^2 > 16$ implies $x^2 > 9$. In particular, $x > 4$ implies $x^2 > 16$. In particular, $x^2 > 16$ implies $x^2 > 9$. Suppose $x > 5$. Then $x > 4$. Hence $x^2 > 16$. So $x^2 > 9$. This proves that if $x > 5$, then $x^2 > 9$.*

Marking scheme. (10 points)

- 2: takes the conjunctive premise apart by two $\wedge$E lines (inside or outside the box).
- 2: opens a box with P (not S) and closes it on S, a line of that box.
- 2: reaches S inside the box of P by $\to$E steps, each citing an implication and exactly its hypothesis, every cited line available.
- 1: each closing $\to$I cites its box as a range.
- 3: prose with one sentence per step, "suppose" for the assumption, and a closing sentence stating the implication.

A route that first derives $Q \to S$ in a box of Q (as in Proposition 1.4.1) and then uses it in the box of P earns the same points: its $\to$E steps in both boxes count for the third line, and its two $\to$I steps for the fourth.

Fatal errors (0 for the step): a box opened with the conclusion S (the converse error); citing a line of the box after it has closed.

Exercise. C4: The assumptions a line depends on, used later

Each line of a derivation depends on the premises and on the assumptions of the boxes it sits in. Take the first derivation in the proof of Proposition 1.4.2 (lines 1–7).

(a) Beside each line, write the set of line numbers of the premises and assumptions it depends on.

(b) For line 5 and for line 6, write the sequent whose premises are the statements on the lines of that set and whose conclusion is the statement of the line. Show that each is derivable, by turning the open assumptions into premises.

(c) *(Challenge.)* Show that, in any derivation, the set of a line cited by an $\wedge$I, $\wedge$E or $\to$E line is contained in the set of the citing line. Then say how the set of a $\to$I line compares with the set of the last line of the box it cites.

Solution.

(a) Line 1: $\{1\}$. Line 2: $\{1,2\}$. Lines 3, 4 and 5: $\{1,2,3\}$. Line 6: $\{1,2\}$. Line 7: $\{1\}$. The premise is in every set; the assumption of line 3 only in the sets of lines 3–5, the lines of its box; the assumption of line 2 in the sets of lines 2–6.

(b) Line 5 gives $(P \wedge Q) \to R,\ P,\ Q \vdash R$:

1	$(P \wedge Q) \to R$	
2	P	
3	Q	
4	$P \wedge Q$	$\wedge$I 2, 3
5	R	$\to$E 1, 4

Line 6 gives $(P \wedge Q) \to R,\ P \vdash Q \to R$; here the box of Q is kept, since line 6 lies outside it:

1	$(P \wedge Q) \to R$	
2	P	
3	$\quad Q$	asm, for $\to$I
4	$\quad P \wedge Q$	$\wedge$I 2, 3
5	$\quad R$	$\to$E 1, 4
6	$Q \to R$	$\to$I 3–5

In each, the lines down to the chosen one, with the assumptions of the boxes still open there made premises, form a derivation.

(c) A line's set consists of the premises and the assumptions of the boxes containing it. Let line k cite line i by $\wedge$I, $\wedge$E or $\to$E. Then line i is available at line k, so by the scope rule of Definition 1.4.2 every box containing line i also contains line k. Hence the assumption of every such box is in the set of line k, and the premises are in both sets: the set of i is contained in the set of k.

For →I: the box's last line is a line of the box itself, so it lies in the cited box and in the boxes around it. The →I line lies in exactly the boxes around the cited box (the condition for a box to be available). So the set of the →I line is the set of the box's last line with one member removed: the line number of the box's assumption. That is the discharge, seen as a set.

Marking scheme. (10 points)

- 3: all seven sets correct (the premise in every set; line 3's assumption only in the sets of lines 3–5).
- 2: line 5's sequent, with a correct derivation.
- 2: line 6's sequent, with the box of Q kept.
- 3: (c) uses the scope rule for the containment, and names the one assumption that a →I line drops.

Fatal errors (0 for the step): putting line 3 in the set of line 6 or line 7; treating an assumption as a premise of the whole derivation.

1.5 Or, Introduced

Sections 1.2 to 1.4 gave the rules for "and" and "implies". With them we can take a conjunction apart, fire an implication on its hypothesis, and prove an implication by opening a box. This section starts on the third connective of the chapter, "or". It gives the rule that **proves** a disjunction, ∨I, and a habit that goes with it: when the goal is a disjunction, commit to one of its sides only when that side is in hand. Near its end it meets a goal that the rules still cannot reach, and that goal is the question the next section answers.

1.5.1 Throwing information away on purpose

Suppose a proof has just shown that a real number x satisfies $x > 0$, and the definition it is checking asks for $x \geq 0$. Every writer takes the step "so $x \geq 0$" without a pause. Look at what the step does. Section 1.1 noted that the symbol $\leq$ hides a disjunction, so $x \geq 0$ is the statement

$$(x > 0) \vee (x = 0).$$

The step turns a statement into a disjunction of which it is one side.

The step loses something. "$x > 0$" tells us that x is not 0; "$x \geq 0$" no longer does. A proof that takes the step throws information away, and it does so on purpose: the definition asks for the weaker statement, and the stronger one is not in the form it asks for.

Remark. **Convention: inequalities written out.** When a rule has to see the disjunction inside $a \leq b$, a derivation writes it out as $(a < b) \vee (a = b)$, and $a \geq b$ as $(a > b) \vee (a = b)$: a checker compares symbols, so ∨I can write "$x \geq 0$" from "$x > 0$", and ∨E can split on it, only in the written-out form. Otherwise $a \leq b$ may stand in a derivation as one statement, and prose keeps the short forms.

The same move appears whenever a known fact has a disjunction as its hypothesis. School algebra says that if $a = 0$ or $b = 0$, then $a \cdot b = 0$. To use this on a product whose first factor we know to be 0, we must hand the fact its hypothesis, "$a = 0$ or $b = 0$". What we have is only

"$a = 0$". So again the step goes from one side to the whole disjunction, and only then does the fact fire.

Both steps are instances of one rule. This section states it as a pattern, reads it in three bullets, and then uses it in two ways: when the goal is a disjunction, and when a premise needs one.

1.5.2 The rule of or-introduction

To prove a disjunction, it is enough to prove one of its two sides.

The rule is defined, like every rule since Section 1.2, by its schema (Definition 1.2.1). The Greek letters stand for arbitrary statements, and i, k for line numbers. The rule has two forms, one for each side, so it is printed as two schemas.

Definition 1.5.1. Or-Introduction

i	φ			i	φ	
k	$\varphi \vee \psi$	$\vee$I i		k	$\psi \vee \varphi$	$\vee$I i

The rule given by these two schemas is called $\vee$I (**or-introduction**).

- *Schema:* the two in the block above, one for each side.
- *Reading:* one available line (an earlier line, not inside a box that has closed), cited alone; write a disjunction with that line's statement as one **whole** side, left or right, and **any** statement as the other side: true or false, related or not.
- *Example:* if line 4 is $x > 0$, then "$\vee$I 4" may write $(x > 0) \vee (x = 0)$, or $(x = 0) \vee (x > 0)$, or the silly but correct $(x > 0) \vee (x < -7)$.

In mathematics "or" is inclusive (Section 1.1): $\varphi \vee \psi$ is true when φ is true, when ψ is true, and when both are. So $\varphi \vee \psi$ is true whenever φ is, whatever ψ says. The rule is this first clause read as a permission.

$\vee$I *loses information, never truth.*

↻ **Second reading** · *Edge cases and near misses*

The degenerate case is ψ equal to φ. From P the rule writes $P \vee P$, and then the two schemas give the same line. It is a legal step, and the next section uses the statement $P \vee P$ once.

The near misses, steps that look like $\vee$I and are not, are listed after the rule card at the end of this section.

Warning. Do not read "or" as "one but not both" when checking a step. From $x = 2$, $\vee$I gives "$x = 2$ or $x \leq 5$", although at $x = 2$ both sides are true. The step is correct: the disjunction is inclusive, and the rule never claims that the added side is false.

Warning. $\vee$I runs one way only. From φ it gives $\varphi \vee \psi$; from $\varphi \vee \psi$ it gives nothing, and in general φ does not follow from $\varphi \vee \psi$ at all. The phrase "in particular" belongs to $\wedge$E, not to "or". For a concrete failure, let x be a real number: "$x \geq 0$ or $x < 0$" is true whatever x is, and "in particular $x \geq 0$" is false at $x = -1$.

Quick check. Let x be a real number, and suppose line 1 of a derivation is $x = 2$. Which of the following may line 2 be, justified by one application of $\vee$I to line 1?

(i) $(x = 2) \vee (x = 3)$.
(ii) $(x = 3) \vee (x = 2)$.
(iii) $(x = 2) \vee (1 = 0)$.
(iv) $x \geq 2$.

Solution.

(i) Yes, by the first schema, with ψ replaced by $x = 3$. (ii) Yes, by the second schema. (iii) Yes, by the first schema, with ψ replaced by the false statement $1 = 0$. The line is true, since its left side is, and it is useless; the rule allows it. (iv) Only once $x \geq 2$ is written out as $(x > 2) \vee (x = 2)$, as the convention on inequalities above asks when a rule must see the disjunction. In that form it is an instance of the second schema, with φ replaced by $x = 2$. A checker who matches symbols sees the difference.

1.5.3 A disjunction as a goal

Section 1.2 asked two questions before writing any line: *what is the goal's main connective?* (goal-directed reasoning, from the bottom up) and *what are the available lines' main connectives?* (premise-directed reasoning, from the top down). Each rule so far added an answer. A conjunction as a goal splits into two subgoals, both needed ($\wedge$I); an implication as a goal opens a box ($\rightarrow$I, Definition 1.4.1).

Now $\vee$ joins the list of goals. If the goal is $\varphi \vee \psi$, its last line can be $\vee$I, and then the subgoal is φ **or** ψ: one of the two, of our choosing. That choice is new, and a wrong choice can leave us working towards a statement that does not follow from what we have. Suppose the premise is $x > 1$ and the goal is $(x > 3) \vee (x > 0)$. Choose the left side, and the new goal $x > 3$ does not follow from the premise: try to break the step, and $x = 2$ does, since there $x > 1$ holds and $x > 3$ fails. Choose the right side, and $x > 0$ follows from $x > 1$.

In the Givens/Goal table, $\vee$I replaces the goal $\varphi \vee \psi$ by the side you choose.

So the goal-directed step for $\vee$ comes with a caution:

Postpone a disjunction: choose a side of a goal $\varphi \vee \psi$ only when that side is in hand.

"In hand" means that the side stands on an available line. Until then, work from the premises: take conjunctions apart, fire implications, open the boxes the goal asks for, and look at what appears. When one side of the goal turns up, the choice has been made for us.

Quick check. Let x be a real number. The premise is $0 < x < 1$ and the goal is $(x < 0) \vee (x < 2)$. Which side should $\vee$I take, and why not the other?

Solution. The right side, $x < 2$: it follows from $x < 1$, the right half of the premise. The left side $x < 0$ does not follow: at $x = \frac{1}{2}$ the premise holds and $x < 0$ fails.

A side chosen too early is worth watching once, all the way to the point where the search stalls, and then once more done right.

Example. A side chosen too early

Show that $P \to R \vdash (P \wedge Q) \to (S \vee R)$.

Solution. The box. The goal is an implication, so we open a box with its hypothesis $P \wedge Q$, and the goal inside is $S \vee R$. The first picture shows the box with its goal at the bottom, and the line that will close it, not yet justified; line 2 records the plan, "asm, for $\to$I".

1	$P \to R$	
2	$\quad P \wedge Q$	asm, for $\to$I
	$\quad \vdots$	
5	$\quad S \vee R$	
6	$(P \wedge Q) \to (S \vee R)$	?

1	$P \to R$	
2	$\quad P \wedge Q$	asm, for $\to$I
	$\quad \vdots$	
4	$\quad S$	
5	$\quad S \vee R$	$\vee$I 4
6	$(P \wedge Q) \to (S \vee R)$	?

The early choice. S is the left side of the goal, so suppose we aim at it at once: line 5 becomes "$S \vee R$, $\vee$I 4", and line 4, S, is the new subgoal (the second picture). Now the search stalls. Nothing mentions S except the goal: no premise, no assumption, no part of either.

Try to break the subgoal. Let x be a real number, and take P: "$x > 1$", Q: "$x > 0$", R: "$x^2 > x$", S: "$x > 5$". The premise $P \to R$ holds for every x: if $x > 1$, then $x^2 = x \cdot x > x$. At $x = 2$ the assumption $P \wedge Q$ is true and S is false. So S does not follow from the lines in hand. The goal was not at fault, since at $x = 2$ its right side R is true ($4 > 2$); the choice was.

Redone. Go back to the choice and postpone it. Work from the lines in hand: $\wedge$E takes P out of the assumption, the premise fires on it, and R, the right side of the goal, is in hand. Only now does $\vee$I choose a side.

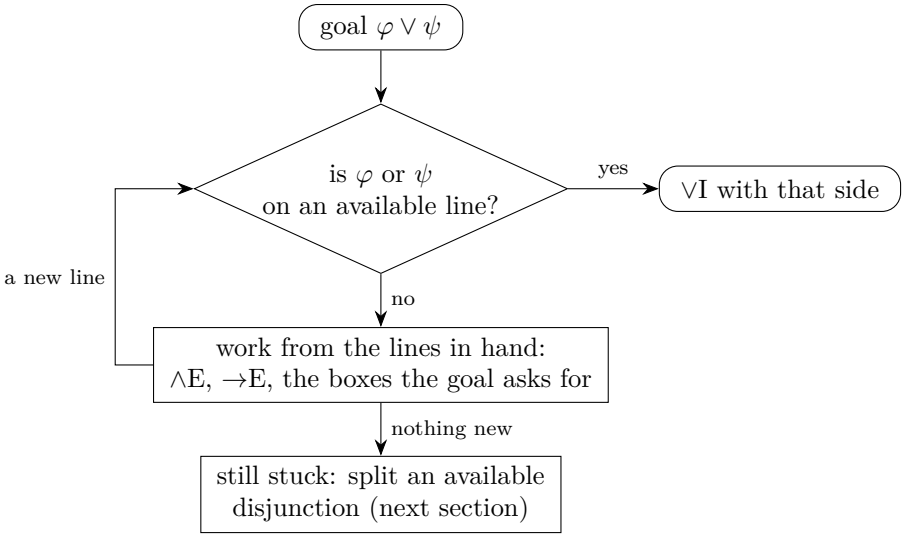
1	$P \to R$	
2	$\quad P \wedge Q$	asm, for $\to$I
3	$\quad P$	$\wedge$E 2
4	$\quad R$	$\to$E 1, 3
5	$\quad S \vee R$	$\vee$I 4
6	$(P \wedge Q) \to (S \vee R)$	$\to$I 2–5

Line 5 uses the second schema, with R on the right. A stalled branch is abandoned, not defended: go back to the last choice and make it later.

Quick check. In Example (A side chosen too early), which side would $\vee$I take if the premise were $P \to S$ instead?

Solution. The left side, S: with the premise $P \to S$, the lines $\wedge$E and $\to$E would put S in hand at line 4. The postponing is the same; the choice is made when a side appears.

The habit of this section as a picture:



Here are four derivations, in order of difficulty. Each Idea says, for each rule, whether it was chosen by the goal or by a premise.

Example. Derivations with or-introduction

Show that:

(a) $P \vdash P \vee Q$, and $P \vdash Q \vee P$;
(b) $P \wedge Q \vdash P \vee Q$;
(c) $P \to Q \vdash P \to (Q \vee R)$;

(d) $(P \wedge Q) \to R \vdash P \to (Q \to (R \vee S))$.

Solution.

(a) *Idea.* The goal is a disjunction, and one of its sides, P, is the premise: in hand from the start. The only choice is which schema, and the goal decides it: P sits on the left of $P \vee Q$ and on the right of $Q \vee P$.

1	P	
2	$P \vee Q$	$\vee$I 1

1	P	
2	$Q \vee P$	$\vee$I 1

In the first, line 2 is the first schema with φ replaced by P and ψ by Q; in the second, it is the second schema with the same replacement. The letter Q appears in no premise. That is the Reading's "**any** statement" at work.

(b) *Idea.* The goal $P \vee Q$ is a disjunction, and neither side stands on a line yet, so we postpone it. Premise-directed: the premise's main connective is $\wedge$, so $\wedge$E takes it apart. After that, **both** sides of the goal are in hand, and either choice finishes.

1	$P \wedge Q$	
2	P	$\wedge$E 1
3	$P \vee Q$	$\vee$I 2

1	$P \wedge Q$	
2	Q	$\wedge$E 1
3	$P \vee Q$	$\vee$I 2

The left derivation chose the left side (first schema); the right one chose the right side (second schema). Both are correct, and the choice was free precisely because both sides were in hand. That will not usually be so.

(c) *Idea.* The goal's main connective is the outer $\to$; the $\vee$ sits inside its conclusion. Goal-directed: open a box with the hypothesis P, and the new goal inside it is $Q \vee R$, a disjunction, so we postpone it. Premise-directed: the premise $P \to Q$ is an implication whose hypothesis P is now on a line, so $\to$E fires and gives Q. The left side of the goal is in hand; $\vee$I takes it, and $\to$I closes the box. The pictures show the box opened with its goal at the bottom, and then the box with the side Q written; only the justification of line 4, marked "?", is left.

1	$P \to Q$	
2	$\quad P$	asm, for $\to$I
	$\quad \vdots$	
4	$\quad Q \vee R$	
5	$P \to (Q \vee R)$	$\to$I 2–4

1	$P \to Q$	
2	$\quad P$	asm, for $\to$I
3	$\quad Q$	$\to$E 1, 2
4	$\quad Q \vee R$	?
5	$P \to (Q \vee R)$	$\to$I 2–4

1	$P \to Q$	
2	$\quad P$	asm, for $\to$I
3	$\quad Q$	$\to$E 1, 2
4	$\quad Q \vee R$	$\vee$I 3
5	$P \to (Q \vee R)$	$\to$I 2–4

Line 3 cites line 1, which lies in no box, and line 2, the assumption of the box that contains line 3. Line 4 is a line of the box itself, so $\to$I may cite the box 2–4.

(d) *Idea.* The goal has two arrows at the top of its structure, $P \to (Q \to \cdots)$, so, as in Proposition 1.4.2, we open two nested boxes, first P, then Q; the inner one will close

first. The innermost goal is $R \vee S$: postpone. Premise-directed: the only line that mentions R is the premise $(P \wedge Q) \to R$. Its hypothesis is $P \wedge Q$, which is not on any line, but both of its halves are: the two assumptions. So $\wedge$I builds it, $\to$E fires, and R is in hand. Then $\vee$I, and the two boxes close. In the pictures, the two boxes are open with the innermost goal at the bottom; then lines 4 and 5 are written, and only the justification of line 6 is left.

1	$(P \wedge Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad Q$	asm, for $\to$I
	$\quad\quad \vdots$	
6	$\quad\quad R \vee S$	
7	$\quad Q \to (R \vee S)$	$\to$I 3–6
8	$P \to (Q \to (R \vee S))$	$\to$I 2–7

1	$(P \wedge Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad Q$	asm, for $\to$I
4	$\quad\quad P \wedge Q$	$\wedge$I 2, 3
5	$\quad\quad R$	$\to$E 1, 4
6	$\quad\quad R \vee S$	?
7	$\quad Q \to (R \vee S)$	$\to$I 3–6
8	$P \to (Q \to (R \vee S))$	$\to$I 2–7

1	$(P \wedge Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad Q$	asm, for $\to$I
4	$\quad\quad P \wedge Q$	$\wedge$I 2, 3
5	$\quad\quad R$	$\to$E 1, 4
6	$\quad\quad R \vee S$	$\vee$I 5
7	$\quad Q \to (R \vee S)$	$\to$I 3–6
8	$P \to (Q \to (R \vee S))$	$\to$I 2–7

This derivation uses every rule of the chapter so far except $\wedge$E and R. Line 4 cites line 2, from the outer box, which is still open at line 4. Line 5 cites line 1 from inside two boxes: a line in no box is available everywhere below it. Line 7 is a line of the outer box itself, which is why $\to$I at line 8 may cite 2–7.

Warning. Choosing the wrong side of a disjunctive goal too early can leave a goal that does not follow. Derivation (c) is where postponing pays. Had we committed at once, inside the box, to the goal's right side, planning to reach $Q \vee R$ from R, the box would need R itself, from the premise $P \to Q$ and the assumption P alone. With x a real number, P: "$x > 2$", Q: "$x > 1$" and R: "$x > 3$", the premise and the assumption hold at $x = 2.5$, and R fails there, so R does not follow. Keep the goal as the whole disjunction until one side appears on a line.

1.5.4 Manufacturing a hypothesis

In every derivation so far, $\vee$I produced the goal, or the goal of a box. It has a second use. A premise whose hypothesis is a disjunction can fire only when that disjunction stands on a line, **exactly** as written (the matching clause of Definition 1.3.1). If one side of it is in hand, $\vee$I builds the whole hypothesis.

Example. A disjunctive hypothesis splits into two implications

Show that $(P \vee Q) \to R \vdash (P \to R) \wedge (Q \to R)$.

Solution. Idea. Goal-directed: the goal is a conjunction, so $\wedge$I splits it into two subgoals, $P \to R$ and $Q \to R$, and each is an implication, so each gets its own box. Inside the box of P, the goal is R. Premise-directed: the only line that mentions R is the premise, and $\to$E on it needs its hypothesis $P \vee Q$ on a line. We have P, one side of that hypothesis. So $\vee$I manufactures $P \vee Q$, and the premise fires. Here $\vee$I is chosen not because the goal is a disjunction (the goal is R) but because a premise's hypothesis is one. The box of Q is the same, with the other schema. In the pictures, both boxes are open with their goals at the bottom; then the $\vee$I lines that manufacture the hypothesis are written, and only the $\to$E lines are left.

1	$(P \vee Q) \to R$	
2	$\quad P$	asm, for $\to$I
	$\quad \vdots$	
4	$\quad R$	
5	$P \to R$	$\to$I 2–4
6	$\quad Q$	asm, for $\to$I
	$\quad \vdots$	
8	$\quad R$	
9	$Q \to R$	$\to$I 6–8
10	$(P \to R) \wedge (Q \to R)$	$\wedge$I 5, 9

1	$(P \vee Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad R$	?
5	$P \to R$	$\to$I 2–4
6	$\quad Q$	asm, for $\to$I
7	$\quad P \vee Q$	$\vee$I 6
8	$\quad R$	?
9	$Q \to R$	$\to$I 6–8
10	$(P \to R) \wedge (Q \to R)$	$\wedge$I 5, 9

1	$(P \vee Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad R$	$\to$E 1, 3
5	$P \to R$	$\to$I 2–4
6	$\quad Q$	asm, for $\to$I
7	$\quad P \vee Q$	$\vee$I 6
8	$\quad R$	$\to$E 1, 7
9	$Q \to R$	$\to$I 6–8
10	$(P \to R) \wedge (Q \to R)$	$\wedge$I 5, 9

Line 3 uses the first schema of Definition 1.5.1 (P on the left), line 7 the second (Q on the right). Both produce the same statement $P \vee Q$, the one the premise asks for. The two boxes are separate: the second opens after the first has closed, and nothing in it cites a line of the first. Line 10 cites lines 5 and 9, which lie in no box.

The product example below, and several exercises, use $\vee$I the same way. Weakening is how a strong fact is fed to a rule, or a theorem, that asks for a weaker one.

In mathematical terms, the derivation says that a theorem with a hypothesis "P or Q" splits into two theorems. Let x be a real number, and suppose we know "if $x \geq 0$, then $x^3 \geq 0$", that is, "if $x > 0$ or $x = 0$, then $x^3 \geq 0$". Then we know "if $x > 0$, then $x^3 \geq 0$" and "if $x = 0$, then $x^3 \geq 0$". The way back, from the two implications to the single one with the disjunctive hypothesis, needs the rule of the next section; it is derived there.

Quick check. In the derivation of Example (A disjunctive hypothesis splits into two implications), line 3 could not have been skipped by letting line 4 cite line 2 directly, as "R — $\to$E 1, 2". Why not?

Solution. →E (Definition 1.3.1) needs the implication on one cited line and its hypothesis, exactly, on the other. The hypothesis of line 1 is $P \vee Q$, and line 2 is P, which does not match it symbol for symbol. Line 3 writes the hypothesis out, and only then does line 1 fire.

1.5.5 A goal we cannot reach yet

The most obvious fact about "or" is that the order of its sides does not matter: $P \vee Q \vdash Q \vee P$. Try to derive it with the rules so far.

The goal $Q \vee P$ is a disjunction, so we postpone it and look at the premise. Its main connective is $\vee$, and no rule so far takes a disjunction apart. $\wedge$E needs a conjunction and →E an implication; $\wedge$I, $\vee$I, and R may cite $P \vee Q$, but only to copy it or to build something larger around it. So the premise gives us nothing, and working from it alone, neither side of the goal ever comes into hand.

We can get this far, though. Each side of the premise, **assumed** in a box, leads to the goal:

1	$P \vee Q$	
2	$\quad P$	asm, for →I
3	$\quad Q \vee P$	$\vee$I 2
4	$P \to (Q \vee P)$	→I 2–3
5	$\quad Q$	asm, for →I
6	$\quad Q \vee P$	$\vee$I 5
7	$Q \to (Q \vee P)$	→I 5–6

This derivation is correct, but it shows $P \vee Q \vdash Q \to (Q \vee P)$, and line 1 is never cited. (A derivation need not use every premise, Definition 1.2.3.) What we would like is to put the three lines 1, 4, and 7 together: *one of P, Q holds, and either one leads to $Q \vee P$; so $Q \vee P$.* That is a proof by cases, and none of the rules so far performs it. With them we cannot find a derivation of $P \vee Q \vdash Q \vee P$.

The missing move is to **use** a disjunction, as $\wedge$E uses a conjunction and →E uses an implication. The next section gives the rule that does this, and with it the six-line derivation that this attempt lacks.

1.5.6 Derive it or break it

From here on, some exercises list claims $\Gamma \vdash \varphi$ without saying which hold. For each, either give a derivation, or break the claim: find statements that make its **counterexample set** $\Gamma, \neg\varphi$ (Definition 1.1.3) true all at once.

For instance, $P \wedge (Q \vee R) \vdash Q$ breaks. Let x be a real number, and take P: "$x > 0$", Q: "$x > 2$", R: "$x < 3$". At $x = 1$ the set $P \wedge (Q \vee R)$, $\neg Q$ is true: $1 > 0$, $1 < 3$, and $1 > 2$ fails.

A claim can also be of neither kind *yet*. $P \vee Q \vdash Q \vee P$ cannot be broken: if $P \vee Q$ is true, one of its sides is, and that side makes $Q \vee P$ true. And with the rules so far no derivation of it has been found: the attempt in "A goal we cannot reach yet" stalls, and the claim waits for the next section's rule. The lists in this section's exercises contain only claims that the rules so

far settle.

1.5.7 How this looks in a mathematical proof

In written mathematics $\vee$I is almost always silent. Here are the two steps of the opening, written out as derivations.

Example. A product with a zero factor

Let a and b be real numbers, and take as known the school fact "if $a = 0$ or $b = 0$, then $a \cdot b = 0$". Prove that if $a = 0$, then $a \cdot b = 0$, naming the rule behind each step. Then write the step from $x > 0$ to $x \geq 0$ as a derivation.

Solution. The goal is an implication, so the proof supposes $a = 0$ (a box). Inside, the fact needs its hypothesis "$a = 0$ or $b = 0$", which $\vee$I manufactures, as in Example (A disjunctive hypothesis splits into two implications). In prose, with each step's rule in brackets:

> Suppose $a = 0$. Then $a = 0$ or $b = 0$ [$\vee$I]. So, by the school fact, $a \cdot b = 0$ [$\to$E]. Hence, if $a = 0$, then $a \cdot b = 0$ [$\to$I].

As a derivation, with the school fact as the premise:

1	$(a = 0 \vee b = 0) \to a \cdot b = 0$	
2	$\quad a = 0$	asm, for $\to$I
3	$\quad a = 0 \vee b = 0$	$\vee$I 2
4	$\quad a \cdot b = 0$	$\to$E 1, 3
5	$a = 0 \to a \cdot b = 0$	$\to$I 2–4

A written proof would say "if $a = 0$, then $a \cdot b = 0$ by the fact", and leave out the words "$a = 0$ or $b = 0$". The derivation shows what that sentence leaves out: a box, a weakening, and one use of the fact.

For the second step, let x be a real number with $x > 0$. Written out, $x \geq 0$ is $(x > 0) \vee (x = 0)$, and the step is one application of the first schema:

1	$x > 0$	
2	$(x > 0) \vee (x = 0)$	$\vee$I 1

Line 2 is $x \geq 0$, written out.

Every "so $x \geq 0$" after $x > 0$, every "so $|t| \leq 1$" after $|t| < 1$, and every "so n is even or $n > 5$" after "n is even" is this rule. A reader who checks a proof line by line should see it, even where the writer did not name it. It is a step that never leads from a truth to a falsehood, and it is also the step that explains why a conclusion can say less than the proof knew.

1.5.8 The rule card

Before the card, the new rule at work on statements larger than single letters. The schema sees only the whole cited line; how much is inside it does not matter.

Example. Instances of ∨I

Schema	φ	ψ	Cited line	New line
first	$P \to Q$	$R \wedge \neg S$	$P \to Q$	$(P \to Q) \vee (R \wedge \neg S)$
second	$P \vee Q$	$P \wedge \neg R$	$P \vee Q$	$(P \wedge \neg R) \vee (P \vee Q)$
second	$(P \wedge Q) \to R$	$\neg R$	$(P \wedge Q) \to R$	$\neg R \vee ((P \wedge Q) \to R)$

Not an instance. From $P \vee Q$, the line $P \vee (Q \vee R)$ is not written by ∨I: the rule adds a side to the **whole** cited line, giving $(P \vee Q) \vee R$ or $R \vee (P \vee Q)$; it never reaches inside.

The card grows by one row. *Shape* shows the given lines (or box), then the new line.

Rule	Shape	You write	Lean
∧I	$\varphi, \psi \Longrightarrow \varphi \wedge \psi$	"We show P, and we show Q."	`constructor`
∧E	$\varphi \wedge \psi \Longrightarrow \varphi$ (or ψ)	"In particular, P."	`h.1` (and `h.2`)
→E	$\varphi \to \psi, \varphi \Longrightarrow \psi$	"Since P implies Q, and P holds, Q."	`exact h hp`
→I	box φ … $\psi \Longrightarrow \varphi \to \psi$	"Assume P. … Hence Q."	`intro hp`
R	φ (available) $\Longrightarrow \varphi$, copied	"As shown above, P."	`exact hp`
∨I	$\varphi \Longrightarrow \varphi \vee \psi$ (or $\psi \vee \varphi$)	"So P or Q holds."	`left` (or `Or.inl hp`); `right`

The phrase is different from the "in particular" of ∧E on purpose. "In particular" announces that a part is being taken out of something bigger; "so P or Q" announces that something bigger, and weaker, is being built.

Warning (Not allowed: ∨I). Each step is offered as one application of ∨I; each breaks the Reading of the rule where shown.

- "P from $P \vee Q$": it runs the rule backwards. The cited line is a side of the new line, never the other way round.
- "$P \wedge Q$ from P": the new line must be a disjunction.
- "$Q \vee P$ from $P \vee Q$": the cited line must be one **whole** side of the new line.
- "$P \vee (Q \vee R)$ from $P \vee Q$": the rule never reaches inside the cited line.

- "$x \geq 0$ from $x > 0$, as written": a checker compares symbols. Write $x \geq 0$ out first, as the convention on inequalities at the start of this section asks.

Quick check. Line 3 is Q. Which of $Q \vee (P \wedge R)$, $(P \wedge R) \vee Q$, $Q \vee Q$, $Q \wedge Q$ may "$\vee$I 3" write?

Solution. The first three. The first is the first schema and the second the second schema, with ψ replaced by $P \wedge R$; the third is either schema, with ψ replaced by Q itself. $Q \wedge Q$ is a conjunction, written by $\wedge$I 3, 3.

Remark (Looking ahead). Chapter 4 proves that a claim that breaks has no derivation, with any of the rules of Chapters 1 and 2, and that a claim that cannot be broken has a derivation with them.

1.5.9 Exercises

1.5.9.1 A. Check your understanding

Exercise. A1

(a) Write the two schemas of Definition 1.5.1 from memory, and say which letter in them may stand for a statement that appears in no earlier line.

(b) Which of the following sentences, taken from proofs about a real number x or real numbers a, b, use $\vee$I? For each of the others, name the rule it uses.
- "We have $x < 3$, so $x \leq 3$."
- "We found that $a = 0$. So $a = 0$ or $b = 0$, and therefore $a \cdot b = 0$."
- "Since $0 < x < 1$, in particular $x < 1$."
- "Since $|x - 2| < 1$ implies $1 < x < 3$, and $|x - 2| < 1$, we get $1 < x < 3$."

(c) True or false, with a reason: $\vee$I may introduce a side that is false.

(d) To show Q, $(P \vee Q) \to (Q \to R) \vdash R$, which rule writes the first line after the premises, and what is that line? Why is that rule used, although the goal is not a disjunction?

Solution.

(a) From a line i holding φ, a line k holding $\varphi \vee \psi$, justified "$\vee$I i"; and, from the same line, a line k holding $\psi \vee \varphi$, with the same justification (Definition 1.5.1). The letter ψ appears in no given line: it may be any statement.

(b) The first uses $\vee$I: $x \leq 3$ is $(x < 3) \vee (x = 3)$, and its left side is in hand. The second uses $\vee$I to build "$a = 0$ or $b = 0$" from its left side, and then $\to$E on the fact about products, as in Example (A product with a zero factor). The third uses $\wedge$E: $0 < x < 1$ is $(0 < x) \wedge (x < 1)$, and the sentence takes out its right half. The fourth uses $\to$E, with the implication first and its hypothesis second.

(c) True. The added side ψ may be any statement, false ones included, as in the line

$(x = 2) \vee (1 = 0)$ of the Quick check after the definition. The new line is still true whenever the cited line is, because "or" is inclusive: one true side is enough.

(d) The rule is $\vee$I, by its second schema, and line 3 is $P \vee Q$, from line 1. It is used because a premise's hypothesis is a disjunction. The goal R occurs only in the premise $(P \vee Q) \to (Q \to R)$, and $\to$E fires on that premise only when $P \vee Q$ itself stands on a line. Only its right side Q is on a line, and $\vee$I turns it into $P \vee Q$. Then $Q \to R$ follows by $\to$E 2, 3, and R by $\to$E 4, 1. The move is the one of Example (A disjunctive hypothesis splits into two implications), where $\vee$I builds a premise's hypothesis.

Exercise. A2: Drill

Each part takes a line or two. The parts run from single steps, through a missing line and a justification column, to whole derivations. Let x be a real number and n an integer.

(a) Line 2 is $(P \to Q) \wedge R$. Write two lines that "$\vee$I 2" can give, one by each schema.

(b) Line 6 is $(R \wedge S) \vee (P \to Q)$, justified "$\vee$I 4". What can line 4 be?

(c) Justify each line: 1 $P \wedge Q$, premise; 2 R, premise; 3 Q; 4 $Q \vee R$; 5 $(Q \vee R) \wedge R$; 6 $S \vee ((Q \vee R) \wedge R)$.

(d) Line 1 is $x = 5$. Which of these may be written as "$\vee$I 1"? (i) $(x = 5) \vee (x = 6)$; (ii) $(x > 4) \vee (x = 5)$; (iii) $x \geq 5$, as written; (iv) $(x = 5) \wedge (x \geq 0)$.

(e) Line 1 is $P \wedge Q$. Is "$P \vee Q$ — $\vee$I 1" legal?

(f) Line 1 is $Q \vee P$. Is "$P \vee Q$ — $\vee$I 1" legal?

(g) Is "$x > 0$, so $x \geq 0$" one step of $\vee$I?

(h) Line 4 is $P \wedge Q$. Is "$(P \vee R) \wedge Q$ — $\vee$I 4" correct? If not, how many lines does it take?

(i) "We have $n = 4$. So n is even or $n > 10$." Which steps does this sentence take?

(j) The premise is $x < -1$ and the goal is $(x < -5) \vee (x < 0)$. Which side should $\vee$I take?

(k) Derive $P \vdash (P \vee Q) \wedge (R \vee P)$.

(l) Derive $P \wedge Q \vdash Q \vee R$.

(m) Derive $P \to Q \vdash P \to (R \vee Q)$.

(n) Derive $(P \vee Q) \to R,\ Q \vdash R$.

(o) Derive $P \to (Q \wedge R) \vdash P \to (S \vee R)$.

Solution.

(a) $((P \to Q) \wedge R) \vee S$ by the first schema and $S \vee ((P \to Q) \wedge R)$ by the second, with any statement for S.

(b) $R \wedge S$ (first schema) or $P \to Q$ (second schema): line 4 is one whole side of line 6.

(c) 3 $\wedge$E 1; 4 $\vee$I 3 (first schema); 5 $\wedge$I 4, 2; 6 $\vee$I 5 (second schema).

(d) (i) Yes, by the first schema. (ii) Yes, by the second schema: $x = 5$ is the right side. (iii) No: as written, $x \geq 5$ is one statement. Written out as $(x > 5) \vee (x = 5)$, it is the second schema. (iv) No: the new line is not a disjunction.

(e) No: $P \wedge Q$ is not a side of $P \vee Q$. From line 1, $\vee$I could write $(P \wedge Q) \vee R$.

(f) No: $Q \vee P$ is not a whole side of $P \vee Q$.

(g) Yes, once $x \geq 0$ is written out as $(x > 0) \vee (x = 0)$: it is the first schema.

(h) No. The new line of $\vee$I is a disjunction with line 4 as one whole side; this line is a conjunction, and it reaches inside line 4. It takes four lines: 5 P, $\wedge$E 4; 6 Q, $\wedge$E 4; 7 $P \vee R$, $\vee$I 5; 8 $(P \vee R) \wedge Q$, $\wedge$I 7, 6.

(i) Two steps: algebra ($4 = 2 \cdot 2$, so n is even), then $\vee$I. From "$n = 4$" itself, $\vee$I gives only "$n = 4$ or ...".

(j) The right side, $x < 0$, which follows from $x < -1$. The left side fails at $x = -2$, where the premise holds and $x < -5$ does not.

(k) 2 $P \vee Q$, $\vee$I 1; 3 $R \vee P$, $\vee$I 1; 4 $(P \vee Q) \wedge (R \vee P)$, $\wedge$I 2, 3.

(l) 2 Q, $\wedge$E 1; 3 $Q \vee R$, $\vee$I 2.

(m) A box opened with P: 2 P, assumption, for $\to$I; 3 Q, $\to$E 1, 2; 4 $R \vee Q$, $\vee$I 3; then 5 $P \to (R \vee Q)$, $\to$I 2–4.

(n) 3 $P \vee Q$, $\vee$I 2; 4 R, $\to$E 1, 3.

(o) A box opened with P: 2 P, assumption, for $\to$I; 3 $Q \wedge R$, $\to$E 1, 2; 4 R, $\wedge$E 3; 5 $S \vee R$, $\vee$I 4; then 6 $P \to (S \vee R)$, $\to$I 2–5.

1.5.9.2 B. Practice

Exercise. B1: Committing to a side

Show that $P \wedge (Q \to R) \vdash Q \to (R \vee S)$, and that $P \vdash (Q \vee P) \vee R$. For each derivation, say at which line you committed to a side of a disjunctive goal, and why that was safe.

Solution. For the first, the goal is an implication, so open a box with Q; the goal inside is $R \vee S$, which we postpone. Premise-directed, $\wedge$E takes out $Q \to R$, and $\to$E fires on the assumption Q.

1	$P \wedge (Q \to R)$	
2	$\quad Q$	asm, for $\to$I
3	$\quad Q \to R$	$\wedge$E 1
4	$\quad R$	$\to$E 3, 2
5	$\quad R \vee S$	$\vee$I 4
6	$Q \to (R \vee S)$	$\to$I 2–5

We committed at line 5, to the left side R, and only after line 4 had put R on a line. The other conjunct of the premise, P, is never used.

For the second, the goal $(Q \vee P) \vee R$ is a disjunction whose left side, $Q \vee P$, is itself a disjunction, and P is one side of that.

1	P	
2	$Q \vee P$	$\vee$I 1
3	$(Q \vee P) \vee R$	$\vee$I 2

Line 2 uses the second schema (P on the right) and line 3 the first, with φ replaced by $Q \vee P$. We committed twice, at lines 2 and 3, and both times the side was in hand: at line 2 the

side P was the premise, and at line 3 the side $Q \vee P$ was line 2. Choosing the side R at the outer level would have left the subgoal R, which does not follow from P: take P true and R false.

Exercise. B2: The last line

A search for $P \to Q,\ Q \to R \vdash P \to (S \vee R)$ has reached the stage below; only the last line is missing. *Name the move, then complete it.*

1	$P \to Q$	
2	$Q \to R$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q$	$\to$E 1, 3
5	$\quad R$	$\to$E 2, 4
6	$\quad S \vee R$	$\vee$I 5
7	**?**	?

Solution. The move is the one read off the goal: an implication goal opens a box with its hypothesis, and the box is closed by $\to$I once its last line is the implication's conclusion. The box opened with P at line 3, and its last line, line 6, is $S \vee R$, the conclusion of the goal. So line 7 closes the box:

1	$P \to Q$	
2	$Q \to R$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q$	$\to$E 1, 3
5	$\quad R$	$\to$E 2, 4
6	$\quad S \vee R$	$\vee$I 5
7	$P \to (S \vee R)$	$\to$I 3–6

Line 7 cites the box as the range 3–6, not its lines one by one. Inside the box, the disjunction $S \vee R$ was postponed until R came into hand at line 5; the side in hand was the right one, so line 6 uses the second schema.

Exercise. B3: A theorem with a weaker hypothesis

Let x be a real number, and take as known the school fact "if $x \geq 1$, then $x^2 \geq x$". Prove that if $x > 1$, then $x^2 \geq x$. Write the proof as a derivation whose premise is the school fact, with $x \geq 1$ written out as a disjunction, and then as two or three sentences of prose.

Solution. Written out, $x \geq 1$ is $(x > 1) \vee (x = 1)$, so the premise is an implication whose hypothesis is a disjunction. The goal is an implication: open a box with $x > 1$. Inside, $\vee$I manufactures the premise's hypothesis, and $\to$E fires.

1	$(x > 1 \vee x = 1) \to x^2 \geq x$	
2	$\quad x > 1$	asm, for $\to$I
3	$\quad x > 1 \vee x = 1$	$\vee$I 2
4	$\quad x^2 \geq x$	$\to$E 1, 3
5	$x > 1 \to x^2 \geq x$	$\to$I 2–4

In prose: *Suppose $x > 1$. Then $x \geq 1$, so $x^2 \geq x$ by the fact. Hence if $x > 1$, then $x^2 \geq x$.* The words "then $x \geq 1$" are line 3. For instance, at $x = 3$ the conclusion reads $9 \geq 3$.

Exercise. B4: Derive or break

For each claim, either derive it with the rules of Sections 1.2–1.5, or break it by statements about a real number x that make its counterexample set true.

(a) $P \vdash P \vee (Q \wedge R)$.
(b) $P \vdash P \wedge (Q \vee R)$.
(c) $P \to R \vdash (P \vee Q) \to R$.
(d) $(P \vee Q) \to R \vdash P \to R$.
(e) $(P \wedge Q) \to R \vdash (Q \wedge P) \to (R \vee S)$.
(f) $(P \to Q) \wedge (P \to R) \vdash P \to (Q \vee R)$.
(g) $P \to (Q \vee R) \vdash P \to Q$.

Solution.

(a) Derive. The side P of the goal is the premise.

1	P	
2	$P \vee (Q \wedge R)$	$\vee$I 1

(b) Break. Take P: "$x > 0$", Q: "$x > 5$", R: "$x < -5$". At $x = 1$, P is true and $Q \vee R$ is false ($1 > 5$ and $1 < -5$ both fail), so $P \wedge (Q \vee R)$ is false. The counterexample set P, $\neg(P \wedge (Q \vee R))$ is true.

(c) Break. Take P: "$x > 3$", Q: "$x < 0$", R: "$x > 1$". The premise $P \to R$ holds for every x. At $x = -1$, $P \vee Q$ is true ($-1 < 0$) and R is false, so $(P \vee Q) \to R$ is false, and the counterexample set $P \to R$, $\neg((P \vee Q) \to R)$ is true.

(d) Derive. Inside the box of P, $\vee$I manufactures the premise's hypothesis, as in Example (A disjunctive hypothesis splits into two implications).

1	$(P \vee Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad R$	$\to$E 1, 3
5	$P \to R$	$\to$I 2–4

(e) Derive. The premise's hypothesis is $P \wedge Q$, and the box gives $Q \wedge P$: take it apart and join it the other way round, then postpone the goal $R \vee S$ until R is in hand.

1	$(P \wedge Q) \to R$	
2	$\quad Q \wedge P$	asm, for $\to$I
3	$\quad Q$	$\wedge$E 2
4	$\quad P$	$\wedge$E 2
5	$\quad P \wedge Q$	$\wedge$I 4, 3
6	$\quad R$	$\to$E 1, 5
7	$\quad R \vee S$	$\vee$I 6
8	$(Q \wedge P) \to (R \vee S)$	$\to$I 2–7

(f) Derive. One half of the premise is enough.

1	$(P \to Q) \wedge (P \to R)$	
2	$\quad P$	asm, for $\to$I
3	$\quad P \to Q$	$\wedge$E 1
4	$\quad Q$	$\to$E 3, 2
5	$\quad Q \vee R$	$\vee$I 4
6	$P \to (Q \vee R)$	$\to$I 2–5

(g) Break. Take P: "$x > 0$", Q: "$x > 5$", R: "$x < 6$". The premise $P \to (Q \vee R)$ holds for every x, since every real number is > 5 or < 6. At $x = 1$, P is true and Q is false, so the counterexample set $P \to (Q \vee R)$, $\neg(P \to Q)$ is true.

Apart from (e), which simply derives, the claims come in pairs. (a) and (b) are look-alikes, with $\vee$ and $\wedge$ traded; (c) and (d) are converses; (g) is a look-alike of (f), with premise and conclusion traded and weakened. In each pair one claim derives and the other breaks.

1.5.9.3 C. Going deeper

Exercise. C1: Which new rules are safe?

Suppose someone proposes new rules for the derivations of this chapter. Let x be a real number.

(a) Take the rule "from $\varphi \vee \psi$, write φ". Starting from the premise "$x \geq 0$ or $x < 0$", which is true for every real x, derive with it a line that is false at $x = -1$.

(b) $\vee$I adds a side ψ that may be any statement at all. Explain why it can never lead from

a true line to a false one.

(c) Is the rule "from φ, write $\varphi \wedge \psi$, for any statement ψ" safe in the same sense? Justify your answer with an instance.

Solution.

(a) With φ replaced by $x \geq 0$ and ψ by $x < 0$, the proposed rule turns the premise into the line $x \geq 0$. At $x = -1$ the premise is true, since $-1 < 0$, and the line $-1 \geq 0$ is false. So the rule can lead from a true line to a false one.

(b) Suppose the cited line φ is true. The new line $\varphi \vee \psi$ (or $\psi \vee \varphi$) is true when at least one of its sides is true, because "or" is inclusive (Section 1.1), and its side φ is true. So the new line is true, whatever ψ is. The truth or falsity of ψ never enters.

(c) No. The new line $\varphi \wedge \psi$ is true only when both sides are, and ψ is arbitrary. With φ replaced by $x = -1$ and ψ by $x \geq 0$, at $x = -1$ the cited line is true and the new line, "$x = -1$ and $x \geq 0$", is false. The contrast with (b) is the whole difference between $\vee$ and $\wedge$: a disjunction needs one true side, so an arbitrary second side is harmless; a conjunction needs both, so it is not. That is why $\wedge$I asks for both sides on lines.

Marking scheme. (10 points)

- 3: (a) applies the proposed rule to the premise to get $x \geq 0$, and checks at $x = -1$ that the premise is true and the new line false.
- 3: (b) argues from the inclusive meaning of "or": one true side makes the disjunction true, whatever the other side is.
- 3: (c) answers no, with an instance in which φ is true and $\varphi \wedge \psi$ is false.
- 1: (c) names the difference: a conjunction needs both sides true.

Fatal errors (0 for the step): in (a), a premise that is false at $x = -1$; in (b), an argument that assumes ψ is true; in (c), "safe" on the grounds that φ is true.

Exercise. C2: Predicting the manufacture

(a) Show that $(P \vee Q) \to (R \wedge S) \vdash P \to S$.

(b) Show that $(P \vee Q) \to R \vdash (P \wedge Q) \to R$.

(c) In each derivation, which rule built the hypothesis of the premise? Could you have predicted, from the premise's main connective and the shape of its hypothesis alone, that this rule would be needed?

Solution.

(a) The goal is an implication: open a box with P; the goal inside is S. The only line mentioning S is the premise, whose conclusion $R \wedge S$ contains it; to use the premise, $\to$E needs $P \vee Q$, built from P by $\vee$I. Then $\wedge$E takes out S.

1	$(P \vee Q) \to (R \wedge S)$	
2	$\quad P$	asm, for $\to$I
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad R \wedge S$	$\to$E 1, 3
5	$\quad S$	$\wedge$E 4
6	$P \to S$	$\to$I 2–5

(b) The goal is an implication: open a box with $P \wedge Q$; the goal inside is R. The premise fires on $P \vee Q$. The assumption gives P by $\wedge$E, and $\vee$I builds $P \vee Q$.

1	$(P \vee Q) \to R$	
2	$\quad P \wedge Q$	asm, for $\to$I
3	$\quad P$	$\wedge$E 2
4	$\quad P \vee Q$	$\vee$I 3
5	$\quad R$	$\to$E 1, 4
6	$(P \wedge Q) \to R$	$\to$I 2–5

Taking Q at line 3 instead, and building $P \vee Q$ by the second schema, works equally well.

(c) In both, $\vee$I built the hypothesis (line 3 in (a), line 4 in (b)), and in both it could be predicted. The premise's main connective is $\to$, so the premise is used by $\to$E, which needs its hypothesis $P \vee Q$ on a line of its own, exactly. That disjunction occurs only inside the premise, as its hypothesis, and no rule takes the hypothesis out of an implication ($\to$E gives the conclusion). So it had to be built, and the only rule of the chapter so far that builds a disjunction is $\vee$I. The one question left was which side would be in hand; in (b) the conjunction $P \wedge Q$ offered both.

1.6 Using Or: Proof by Cases

The previous section met a goal the rules could not reach: from $P \vee Q$, derive $Q \vee P$. The rules could build a disjunction ($\vee$I, Definition 1.5.1), but none of them could use one. This section gives the rule that uses a disjunction, $\vee$E. It is the most familiar move in mathematics after "suppose": the split into cases, "if $x \geq 0$, then ...; if $x < 0$, then ...". It is also the last rule of this chapter. We use it to reach the stalled goal, to prove the laws that let "and" and "or" be rearranged, and to prove three facts about absolute values; the section ends with all the rules of the chapter on one page.

1.6.1 A proof that needs two cases

The absolute value of a real number is defined in two cases:

$$|x| := \begin{cases} x & \text{if } x \geq 0, \\ -x & \text{if } x < 0. \end{cases}$$

Suppose we want to prove that $|xy| = |x||y|$ for all real numbers x and y. There is no single computation to do. Which formula gives $|xy|$ depends on the sign of xy, and that depends on

the signs of x and y. Yet once the signs are fixed, each case is one line: if $x > 0$ and $y < 0$, for instance, then $|xy| = -xy = x \cdot (-y) = |x||y|$.

So the proof we want has the shape "every real number is ≥ 0 or < 0; in the first case ...; in the second case ...; so in every case ...". The rules so far cannot say this. The attempt at $P \vee Q \vdash Q \vee P$ in Section 1.5 met the same wall from the other side. It derived that each side of the premise, assumed in a box, leads to the goal: lines 4 and 7 there were $P \to (Q \vee P)$ and $Q \to (Q \vee P)$. What it lacked was a rule to say "one of P, Q holds, and either one leads to $Q \vee P$; so $Q \vee P$": a rule that takes "φ or ψ", which does not say **which** side holds, and still reaches a definite conclusion. That is the job of the elimination rule for "or".

1.6.2 The rule of or-elimination

To use "φ or ψ", reach your goal from φ, and reach it again from ψ.

The definition block states both rules for disjunction, so that the pair stands in one place: the new rule first, then the two schemas of $\vee$I recalled from Section 1.5. As in every schema (Definition 1.2.1), the Greek letters stand for arbitrary statements and i, j, k, l, m, n for line numbers.

Definition 1.6.1. Rules for Disjunction

i	$\varphi \vee \psi$	
j	$\quad \varphi$	asm, for $\vee$E
	$\quad \vdots$	
k	$\quad \chi$	
l	$\quad \psi$	asm, for $\vee$E
	$\quad \vdots$	
m	$\quad \chi$	
n	χ	$\vee$E i, j–k, l–m

i	φ	
k	$\varphi \vee \psi$	$\vee$I i

i	φ	
k	$\psi \vee \varphi$	$\vee$I i

The first schema is the rule $\vee$E (**or-elimination**, or **proof by cases**); the other two are $\vee$I, as in Definition 1.5.1. In $\vee$E, each line χ (line k and line m) is a line of its box itself, not of a box inside it.

This rule has conditions that matter; read each one.

- **The disjunction.** Line i holds a disjunction $\varphi \vee \psi$, available at line n. The cases are not ours to invent: they are the two sides of a disjunction we already have, as a premise, an assumption, or a line derived earlier. Writing "Case 1 ... Case 2 ..." does not create it.
- **Two boxes.** The first box opens with the left side φ as its assumption, the second with the right side ψ. They are separate: the second opens after the first has closed, and the drawing shows a break in the vertical bar between them. So, by the scope rule of Definition 1.4.2, no line of the second box may cite a line of the first. The disjunction does not tell us which side holds, so we must be ready for both, and each case's assumption is in force only inside its own box. Each case box's assumption is printed "asm, for $\vee$E"; read it as "Case 1: ..." and "Case 2: ...".
- **The same last line.** Both boxes end in the **same** statement χ, and in each box χ is a

line of that box itself, not of a box opened inside it. The rule then writes χ a third time, outside both boxes, where neither φ nor ψ is assumed. That is honest only because χ was reached whichever case holds.

- **Three citations.** The justification "$\vee$E i, j–k, l–m" cites the disjunction first, then the box of the left side, then the box of the right side, each box as a range. Both boxes must be available at line n: closed, and sitting directly in the part of the derivation where line n is (Definition 1.4.2).
- **Any conclusion.** The statement χ need not mention φ or ψ. It is whatever we are trying to reach. The next subsection says how to choose it.

↻ **Second reading** · *A box of one line*

The degenerate case is a box of one line, whose assumption is already χ. Nothing in the rule asks a box to contain more than its assumption, just as nothing in $\rightarrow$I did (Example (First boxes)). The second derivation below has two such boxes.

Warning. A case may use only its own assumption. Here is the error in its usual form. Let x be a real number, and split on "$x \geq 0$ or $x < 0$". *Case $x \geq 0$: then $|x| = x$. Case $x < 0$: since $x \geq 0$, again $|x| = x$. So $|x| = x$.* The second case cites the assumption of the first, whose box has closed, which the scope rule forbids. The conclusion is false at $x = -1$.

1.6.3 Cases end in the current goal

Section 1.2 set up two questions: *what is the goal's main connective?* and *what are the available lines' main connectives?* Section 1.5 answered the first for $\vee$: postpone a disjunctive goal until one side is in hand. Now we can answer the second. A disjunction on an available line invites $\vee$E: split on it.

What should the two boxes end in? The rule leaves χ to us, and the choice is always the same. It is the **current goal**: whatever we are trying to reach at the point where the split is made. That may be the goal of the whole derivation, or a subgoal inside other boxes. The boxes of the split end in it, and the $\vee$E line that closes the split then writes it outside both boxes, exactly where it was wanted.

Cases end in the current goal.

The two habits fit together. Faced with a goal $\varphi \vee \psi$ and an available disjunction $\alpha \vee \beta$, do not choose a side of the goal yet: split on $\alpha \vee \beta$ first. Inside each case the current goal is still $\varphi \vee \psi$, but now a new line is in hand, the case's assumption, and it may be one side of the goal. Then choose, and the choice may be different in the two boxes. This is exactly how the stalled goal of Section 1.5 goes.

In the Givens/Goal table, each case of a split adds its assumption to the Givens and keeps the same Goal. For the stalled goal:

Givens	Goal
$P \vee Q$	$Q \vee P$

Givens	Goal
$P \vee Q$, P (first case)	$Q \vee P$
$P \vee Q$, Q (second case)	$Q \vee P$

Example. First derivations with "or"

Show that $P \vee Q \vdash Q \vee P$, and that $P \vee P \vdash P$.

Solution. Idea. For the first: the goal $Q \vee P$ is a disjunction, and neither side is in hand, so we postpone it. The premise is a disjunction, so we split on it; the current goal is $Q \vee P$, so both boxes will end there. In the box of P, the right side of the goal is in hand, and $\vee$I takes it; in the box of Q, the left side is, and $\vee$I takes that. The choice of side differs between the boxes, and each is safe because it is made inside its own case. The pictures show the two boxes opened, each ending in the goal; then each box filled, with only line 6 left to justify.

1	$P \vee Q$	
2	$\quad P$	asm, for $\vee$E
	$\quad \vdots$	
3	$\quad Q \vee P$	
4	$\quad Q$	asm, for $\vee$E
	$\quad \vdots$	
5	$\quad Q \vee P$	
6	$Q \vee P$	?

1	$P \vee Q$	
2	$\quad P$	asm, for $\vee$E
3	$\quad Q \vee P$	$\vee$I 2
4	$\quad Q$	asm, for $\vee$E
5	$\quad Q \vee P$	$\vee$I 4
6	$Q \vee P$	?

1	$P \vee Q$	
2	$\quad P$	asm, for $\vee$E
3	$\quad Q \vee P$	$\vee$I 2
4	$\quad Q$	asm, for $\vee$E
5	$\quad Q \vee P$	$\vee$I 4
6	$Q \vee P$	$\vee$E 1, 2–3, 4–5

The two boxes are separate: the break in the vertical bar between lines 3 and 4 shows where the first closes and the second opens. Line 3 uses the second schema of $\vee$I (P on the right) and line 5 the first (Q on the left). Both boxes end in $Q \vee P$, and line 6 cites the disjunction on line 1 and the boxes 2–3 and 4–5. Compare the stalled attempt of Section 1.5: its two boxes (2–3 and 5–6 there) are these two boxes, and line 6 here is the step it could not take.

For the second, the goal is P, and the premise is a disjunction, so we split. Both cases are already the goal, so each box consists of its assumption alone.

1	$P \vee P$	
2	$\quad P$	asm, for $\vee$E
3	$\quad P$	asm, for $\vee$E
4	P	$\vee$E 1, 2–2, 3–3

This is the degenerate case split, in which the two cases are the same case. The box 2–2 is a legal box whose last line is P, a line of that box itself, and so is 3–3.

↻ **Second reading** · *One line short*

Change one line of the first derivation and it stops being a derivation. Suppose the second box stops at its assumption Q, and the last line still claims $Q \vee P$:

1	$P \vee Q$	
2	$\quad P$	asm, for $\vee$E
3	$\quad Q \vee P$	$\vee$I 2
4	$\quad Q$	asm, for $\vee$E
5	$Q \vee P$	$\vee$E 1, 2–3, 4–4

Everything cited exists and is available: the disjunction, a box opened by P, a box opened by Q. What fails is the clause about the last lines. The box 4–4 ends in Q, the box 2–3 ends in $Q \vee P$, and $\vee$E needs one statement that ends both. The repair is the missing line 5 of the correct derivation.

Quick check. A derivation contains a line $P \vee Q$ and a box opened by P that ends in R. No box is opened by Q. Which clause of $\vee$E stops us writing R? What single change to the derivation would allow it?

Solution. The clause "two boxes": the rule needs one box opened by each side of the disjunction, and the box opened by Q is missing. Adding a box opened by Q whose last line is R (a derivation of R from Q and the lines available there) allows $\vee$E, citing the disjunction and both boxes.

↻ **Second reading** · *Why one common last line loses nothing*

Why is the rule built like this? Drop the second box, and from "$x \geq 0$ or $x < 0$" and a derivation of $|x| = x$ from $x \geq 0$ we could conclude $|x| = x$ for every real x, which fails at $x = -1$. Allow the two boxes to end in different statements χ_1 and χ_2, and ask what may be written after them: all we honestly know is "χ_1 or χ_2". That is already reachable with the rule as it stands, by $\vee$I at the end of each box, turning both last lines into $\chi_1 \vee \chi_2$ (exercise B3

carries this out). So the rule asks for one common last line and loses nothing.

The next derivation splits on a disjunction that is not a premise but an **assumption**, opened because the goal asked for it.

Example. Cases on a hypothesis are two implications

Show that $(P \to R) \land (Q \to R) \vdash (P \lor Q) \to R$.

Solution. Idea. Goal-directed: the goal is an implication, so we open a box with its hypothesis $P \lor Q$, and the current goal inside is R. Premise-directed: the assumption just made is a disjunction, so we split on it, and both case boxes end in the current goal R. In the case P, the premise's main connective is $\land$; $\land$E takes out $P \to R$, and $\to$E fires on the case's assumption. The case Q is the same with the other half of the premise. The pictures show the two case boxes opened, each ending in R; then the first case filled in.

1	$(P \to R) \land (Q \to R)$	
2	$\quad P \lor Q$	asm, for $\to$I
3	$\quad\quad P$	asm, for $\lor$E
	$\quad\quad \vdots$	
5	$\quad\quad R$	
6	$\quad\quad Q$	asm, for $\lor$E
	$\quad\quad \vdots$	
8	$\quad\quad R$	
9	$\quad R$	$\lor$E 2, 3–5, 6–8
10	$(P \lor Q) \to R$	$\to$I 2–9

1	$(P \to R) \land (Q \to R)$	
2	$\quad P \lor Q$	asm, for $\to$I
3	$\quad\quad P$	asm, for $\lor$E
4	$\quad\quad P \to R$	$\land$E 1
5	$\quad\quad R$	$\to$E 4, 3
6	$\quad\quad Q$	asm, for $\lor$E
	$\quad\quad \vdots$	
8	$\quad\quad R$	
9	$\quad R$	$\lor$E 2, 3–5, 6–8
10	$(P \lor Q) \to R$	$\to$I 2–9

1	$(P \to R) \wedge (Q \to R)$	
2	$\quad P \vee Q$	asm, for $\to$I
3	$\quad\quad P$	asm, for $\vee$E
4	$\quad\quad P \to R$	$\wedge$E 1
5	$\quad\quad R$	$\to$E 4, 3
6	$\quad\quad Q$	asm, for $\vee$E
7	$\quad\quad Q \to R$	$\wedge$E 1
8	$\quad\quad R$	$\to$E 7, 6
9	$\quad R$	$\vee$E 2, 3–5, 6–8
10	$(P \vee Q) \to R$	$\to$I 2–9

Line 9 closes the split inside the outer box. It is a line of the outer box itself, so $\to$I at line 10 may cite 2–9. Lines 4 and 7 cite line 1, which lies in no box.

Quick check. In the derivation of Example (Cases on a hypothesis are two implications), which lines and boxes does the $\vee$E line cite?

Solution. Line 9 cites the disjunction on line 2 and the two case boxes, as ranges: 3–5 and 6–8. All three sit directly in the outer box 2–9, as line 9 does.

Section 1.5 derived the converse, $(P \vee Q) \to R \vdash (P \to R) \wedge (Q \to R)$, in Example (A disjunctive hypothesis splits into two implications), using $\vee$I and no $\vee$E. So each of the two statements is derivable from the other: *cases on a hypothesis are two implications.* This is how mathematicians prove "if P or Q, then R": they prove "if P, then R" and "if Q, then R". To show that an integer which is one more, or three more, than a multiple of 4 is odd, one treats the two forms separately, and the derivation above is the reason that is enough.

1.6.4 Two cases, two conclusions

When the two cases lead to different statements, each case can still end in the current goal, if the goal is the disjunction of the two.

Example. Constructive dilemma

Show that $(P \to R) \wedge (Q \to S),\ P \vee Q \vdash R \vee S$.

Solution. Idea. The goal $R \vee S$ is a disjunction with no side in hand: postpone. The premise $P \vee Q$ is a disjunction: split, both cases ending in $R \vee S$. In each case $\wedge$E takes out the implication that fits the case, $\to$E fires on the case's assumption, and $\vee$I weakens the result to the goal, on the left in one case and on the right in the other. The pictures show the two case boxes opened, each ending in the goal; then the first box filled.

1	$(P \to R) \wedge (Q \to S)$	
2	$P \vee Q$	
3	$\quad P$	asm, for $\vee$E
	$\quad \vdots$	
6	$\quad R \vee S$	
7	$\quad Q$	asm, for $\vee$E
	$\quad \vdots$	
10	$\quad R \vee S$	
11	$R \vee S$	?

1	$(P \to R) \wedge (Q \to S)$	
2	$P \vee Q$	
3	$\quad P$	asm, for $\vee$E
4	$\quad P \to R$	$\wedge$E 1
5	$\quad R$	$\to$E 4, 3
6	$\quad R \vee S$	$\vee$I 5
7	$\quad Q$	asm, for $\vee$E
	$\quad \vdots$	
10	$\quad R \vee S$	
11	$R \vee S$	?

1	$(P \to R) \wedge (Q \to S)$	
2	$P \vee Q$	
3	$\quad P$	asm, for $\vee$E
4	$\quad P \to R$	$\wedge$E 1
5	$\quad R$	$\to$E 4, 3
6	$\quad R \vee S$	$\vee$I 5
7	$\quad Q$	asm, for $\vee$E
8	$\quad Q \to S$	$\wedge$E 1
9	$\quad S$	$\to$E 8, 7
10	$\quad R \vee S$	$\vee$I 9
11	$R \vee S$	$\vee$E 2, 3–6, 7–10

With numbers: let x be a real number with $x \geq 1$ or $x \leq -1$. Since $x \geq 1$ gives $x^3 \geq 1$, and $x \leq -1$ gives $x^3 \leq -1$, the derivation shows that $x^3 \geq 1$ or $x^3 \leq -1$.

The pattern is called the **constructive dilemma**. When R and S are the same statement there is nothing to weaken: the skeleton of "From a skeleton to a proof about numbers" below reaches R itself.

1.6.5 Laws of "or"

Case splits are not only about numbers. The same two rules prove the laws that let us rearrange "and" and "or". The first says that "and" distributes over "or", in both directions.

Proposition 1.6.1. Distributivity

$P \wedge (Q \vee R) \vdash (P \wedge Q) \vee (P \wedge R)$, and $(P \wedge Q) \vee (P \wedge R) \vdash P \wedge (Q \vee R)$.

Try it first (ten minutes): derive the first sequent. Which rule will give its last line, and which disjunction will you split on?

Idea. The natural first attempt reads the goal: it is a disjunction, so end with $\vee$I and aim at one side, say $P \wedge Q$. That needs Q on a line, and the premise offers only P and $Q \vee R$; nothing turns $Q \vee R$ into Q. The other side fails the same way. So the goal must be postponed, and the attempt shows why: the premise holds a disjunction, $Q \vee R$, and neither of its sides is known. Premise-directed, then: $\wedge$E takes the premise apart into P and $Q \vee R$, and we split on $Q \vee R$, both cases ending in the current goal $(P \wedge Q) \vee (P \wedge R)$. In the case Q, the side $P \wedge Q$ is one $\wedge$I away, and $\vee$I weakens it to the whole goal; the case R builds the other side. In the second sequent the premise itself is the disjunction, so we split on it at once; each case is a conjunction containing P, and $\vee$I turns its other half into $Q \vee R$.

For the first derivation: the two case boxes opened, each ending in the goal; then lines 5 and 8 written, and only lines 6 and 9 left to justify.

1	$P \wedge (Q \vee R)$	
2	P	$\wedge$E 1
3	$Q \vee R$	$\wedge$E 1
4	$\quad Q$	asm, for $\vee$E
	$\quad \vdots$	
6	$\quad (P \wedge Q) \vee (P \wedge R)$	
7	$\quad R$	asm, for $\vee$E
	$\quad \vdots$	
9	$\quad (P \wedge Q) \vee (P \wedge R)$	
10	$(P \wedge Q) \vee (P \wedge R)$	$\vee$E 3, 4–6, 7–9

1	$P \wedge (Q \vee R)$	
2	P	$\wedge$E 1
3	$Q \vee R$	$\wedge$E 1
4	$\quad Q$	asm, for $\vee$E
5	$\quad P \wedge Q$	$\wedge$I 2, 4
6	$\quad (P \wedge Q) \vee (P \wedge R)$	?
7	$\quad R$	asm, for $\vee$E
8	$\quad P \wedge R$	$\wedge$I 2, 7
9	$\quad (P \wedge Q) \vee (P \wedge R)$	?
10	$(P \wedge Q) \vee (P \wedge R)$	$\vee$E 3, 4–6, 7–9

Before reading on: in the first derivation, which line does each box end on, and why must the two be the same?

Proof. For the first sequent, the following is a derivation with premise $P \wedge (Q \vee R)$ and last line $(P \wedge Q) \vee (P \wedge R)$.

1	$P \wedge (Q \vee R)$	
2	P	$\wedge$E 1
3	$Q \vee R$	$\wedge$E 1
4	$\quad Q$	asm, for $\vee$E
5	$\quad P \wedge Q$	$\wedge$I 2, 4
6	$\quad (P \wedge Q) \vee (P \wedge R)$	$\vee$I 5
7	$\quad R$	asm, for $\vee$E
8	$\quad P \wedge R$	$\wedge$I 2, 7
9	$\quad (P \wedge Q) \vee (P \wedge R)$	$\vee$I 8
10	$(P \wedge Q) \vee (P \wedge R)$	$\vee$E 3, 4–6, 7–9

Lines 5 and 8 may cite line 2, which lies in no box and so is available everywhere below it. Both boxes end in the goal, lines 6 and 9, as $\vee$E at line 10 requires.

For the second sequent, the following is a derivation with premise $(P \wedge Q) \vee (P \wedge R)$ and last line $P \wedge (Q \vee R)$.

1	$(P \wedge Q) \vee (P \wedge R)$	
2	$\quad P \wedge Q$	asm, for $\vee$E
3	$\quad P$	$\wedge$E 2
4	$\quad Q$	$\wedge$E 2
5	$\quad Q \vee R$	$\vee$I 4
6	$\quad P \wedge (Q \vee R)$	$\wedge$I 3, 5
7	$\quad P \wedge R$	asm, for $\vee$E
8	$\quad P$	$\wedge$E 7
9	$\quad R$	$\wedge$E 7
10	$\quad Q \vee R$	$\vee$I 9
11	$\quad P \wedge (Q \vee R)$	$\wedge$I 8, 10
12	$P \wedge (Q \vee R)$	$\vee$E 1, 2–6, 7–11

This proves both sequents.

■

The answer to the prompt: the boxes end on lines 6 and 9, both $(P \wedge Q) \vee (P \wedge R)$. They must be the same statement because line 10 writes that statement outside both boxes, where neither Q nor R is assumed, and that is honest only if it was reached in each case.

In the second derivation, line 8 derives P a second time instead of citing line 3. It has to: line 3 lies in the box 2–6, which has closed, so line 3 is not available at line 8. Each case must earn its own P.

↻ **Second reading** · *Distributivity in words*

For a real number x, take P: "$x \neq 0$", Q: "$x > 0$" and R: "$x < 0$". The first sequent then reads: a non-zero number that is positive or negative is a positive non-zero number or a negative non-zero number. The law is dull in words, which is as it should be. Its use is in derivations, where it moves a case split from inside a conjunction to the outside, and in the laws of sets, where "and" and "or" become intersection and union.

The next derivation shows a case split **inside** a case: $\vee$E applied within one box of another $\vee$E. Before the lines, the shape: each split ends with its own $\vee$E line, line 11 for the inner split and line 12 for the outer.

1	$P \vee (Q \vee R)$	
2	$\quad P$	asm, for $\vee$E
	$\quad \vdots$	
4	$\quad (P \vee Q) \vee R$	
5	$\quad Q \vee R$	asm, for $\vee$E
6	$\quad\quad Q$	asm, for $\vee$E
	$\quad\quad \vdots$	
8	$\quad\quad (P \vee Q) \vee R$	
9	$\quad\quad R$	asm, for $\vee$E
	$\quad\quad \vdots$	
10	$\quad\quad (P \vee Q) \vee R$	
11	$\quad (P \vee Q) \vee R$	$\vee$E
12	$(P \vee Q) \vee R$	$\vee$E

Quick check. In the outline, which line closes the inner split, and which line and boxes does it cite?

Solution. Line 11. It cites the disjunction on line 5 and the boxes 6–8 and 9–10, all of which sit directly in the box 5–11.

Example. Regrouping a disjunction

Show that $P \vee (Q \vee R) \vdash (P \vee Q) \vee R$.

Solution. Idea. The goal is a disjunction, and no side of it is in hand: postpone. The premise is a disjunction: split on it, with both boxes ending in the current goal $(P \vee Q) \vee R$. The case P is two weakenings: first to $P \vee Q$, then to the goal. The case $Q \vee R$ is itself a disjunction, and neither Q nor R is known; so, inside that box, we split again, and the inner boxes end in the same current goal. In the inner case Q the goal is two weakenings away, and in the inner case R one. The outline above was the first picture; in the second, every box is filled, and only the two $\vee$E citations are left.

1	$P \vee (Q \vee R)$	
2	$\quad P$	asm, for $\vee$E
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad (P \vee Q) \vee R$	$\vee$I 3
5	$\quad Q \vee R$	asm, for $\vee$E
6	$\quad\quad Q$	asm, for $\vee$E
7	$\quad\quad P \vee Q$	$\vee$I 6
8	$\quad\quad (P \vee Q) \vee R$	$\vee$I 7
9	$\quad\quad R$	asm, for $\vee$E
10	$\quad\quad (P \vee Q) \vee R$	$\vee$I 9
11	$\quad (P \vee Q) \vee R$	$\vee$E
12	$(P \vee Q) \vee R$	$\vee$E

Before reading on: which disjunction does the inner $\vee$E *cite, and which box must contain it?*

1	$P \vee (Q \vee R)$	
2	$\quad P$	asm, for $\vee$E
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad (P \vee Q) \vee R$	$\vee$I 3
5	$\quad Q \vee R$	asm, for $\vee$E
6	$\quad\quad Q$	asm, for $\vee$E
7	$\quad\quad P \vee Q$	$\vee$I 6
8	$\quad\quad (P \vee Q) \vee R$	$\vee$I 7
9	$\quad\quad R$	asm, for $\vee$E
10	$\quad\quad (P \vee Q) \vee R$	$\vee$I 9
11	$\quad (P \vee Q) \vee R$	$\vee$E 5, 6–8, 9–10
12	$(P \vee Q) \vee R$	$\vee$E 1, 2–4, 5–11

Line 11 is the inner case split. It cites the disjunction on line 5, the assumption of the outer box it sits in, and the boxes 6–8 and 9–10, all of which sit inside the box 5–11, as line 11 does. Line 12 is the outer case split. Line 11 is the inner split's own $\vee$E line, so the box 5–11 ends on a line of its own, as Definition 1.6.1 requires.

1.6.6 From a skeleton to a proof about numbers

Almost every proof by cases in analysis has the same skeleton. A known fact supplies the cases, and each case is handled by a separate argument, a small lemma that turns the case's assumption into the goal. As a derivation, with the fact and the two lemmas as premises:

1	$P \vee Q$	
2	$P \to R$	
3	$Q \to R$	
4	$\quad P$	asm, for $\vee$E
5	$\quad R$	$\to$E 2, 4
6	$\quad Q$	asm, for $\vee$E
7	$\quad R$	$\to$E 3, 6
8	R	$\vee$E 1, 4–5, 6–7

The idea behind it is the whole method of this section at its simplest. The goal R is atomic, so the goal suggests nothing. Premise-directed, the disjunction on line 1 is split, both boxes ending in the current goal R. In each box the matching lemma fires on the case's assumption. (Example (Cases on a hypothesis are two implications) did the same with the two lemmas packed into one conjunction, and inside a box.)

In a proof about real numbers, the disjunction is usually a fact about the order. The facts about $\leq$ and the absolute value used below are school algebra: every real number a satisfies $a \geq 0$ or $a < 0$; the rules of signs for products; adding two inequalities of the same direction. Here we take them as known. In the language of this chapter, "every real number a satisfies $a \geq 0$ or $a < 0$" is a supply of disjunctions: for each number we meet, it gives a line $(a \geq 0) \vee (a < 0)$ that $\vee$E may cite.

The first example is the smallest, and it is written twice: as a proof in which every step carries its rule in brackets, and as a derivation.

Example. A number is at most its absolute value

Let x be a real number. Prove that $|x| \geq x$, by cases, labeling each step with its rule.

Solution. The goal $|x| \geq x$ is the disjunction $(|x| > x) \vee (|x| = x)$, and neither side is known for every x (the first fails at $x = 1$, the second at $x = -1$), so we postpone it. The fact "$x \geq 0$ or $x < 0$" is a disjunction to split on. In prose:

> We have $x \geq 0$ or $x < 0$, and we argue by cases [$\vee$E].
>
> *Case* $x \geq 0$. By the definition of the absolute value, $|x| = x$ [$\to$E]. So $|x| \geq x$ [$\vee$I].
>
> *Case* $x < 0$. By the definition, $|x| = -x$, and $-x > 0 > x$; so $|x| > x$ [$\to$E], which is the school-algebra fact "if $x < 0$, then $|x| > x$" used on the case's assumption. So $|x| \geq x$ [$\vee$I].
>
> In either case $|x| \geq x$, so $|x| \geq x$.

As a derivation, with the order fact, the definition's first clause, and the school-algebra fact about negative numbers as premises, and with $|x| \geq x$ written out:

1	$x \geq 0 \vee x < 0$	
2	$x \geq 0 \to \lvert x\rvert = x$	
3	$x < 0 \to \lvert x\rvert > x$	
4	$\quad x \geq 0$	asm, for $\vee$E
5	$\quad \lvert x\rvert = x$	$\to$E 2, 4
6	$\quad \lvert x\rvert > x \vee \lvert x\rvert = x$	$\vee$I 5
7	$\quad x < 0$	asm, for $\vee$E
8	$\quad \lvert x\rvert > x$	$\to$E 3, 7
9	$\quad \lvert x\rvert > x \vee \lvert x\rvert = x$	$\vee$I 8
10	$\lvert x\rvert > x \vee \lvert x\rvert = x$	$\vee$E 1, 4–6, 7–9

It is the skeleton above, with a weakening at the end of each box: line 6 takes the right side of the goal (second schema of $\vee$I) and line 9 the left side (first schema). Each case took the side it had in hand. At $x = 3$, the first case applies and line 6 reads $3 > 3$ or $3 = 3$; at $x = -2$, the second applies and line 8 reads $2 > -2$.

Quick check. In the derivation of Example (A number is at most its absolute value), which statement do lines 6 and 9 both write, and why must they agree?

Solution. Both write $(|x| > x) \vee (|x| = x)$, the current goal written out. They must agree because $\vee$E closes a split only when both case boxes end in the same statement.

The next two examples are longer, and in each the case split is doing real work.

Example. The absolute value of a product

Let x and y be real numbers. Prove that $|xy| = |x||y|$.

Solution. We use that $|a| = -a$ whenever $a \leq 0$: for $a < 0$ this is the definition, and for $a = 0$ both sides are 0.

The number x satisfies $x \geq 0$ or $x < 0$, and we argue by cases [$\vee$E]. Inside each case, $y \geq 0$ or $y < 0$, and we argue by cases again.

Case $x \geq 0$, $y \geq 0$. Then $xy \geq 0$, so $|xy| = xy = |x||y|$.

Case $x \geq 0$, $y < 0$. Then $xy \leq 0$, so $|xy| = -xy = x \cdot (-y) = |x||y|$.

Case $x < 0$, $y \geq 0$. Then $xy \leq 0$, so $|xy| = -xy = (-x) \cdot y = |x||y|$.

Case $x < 0$, $y < 0$. Then $xy > 0$, so $|xy| = xy = (-x)(-y) = |x||y|$.

Each inner case split reaches $|xy| = |x||y|$, so each outer case does, and hence the equation

holds.

The structure is two uses of $\vee$E, one inside each box of the other: two outer boxes (the sign of x), each containing two inner boxes (the sign of y), and every one of the four inner boxes ending in the same equation, the current goal. The second and third cases needed the remark about $a \leq 0$. When $x = 0$ and $y < 0$, the product is 0, not negative, and without the remark the case would have a hole at exactly that point.

A case split need not be the obvious one. The triangle inequality has an equally short proof with two cases, if we split on the sign of the right quantity.

Example. The triangle inequality by cases

Let x and y be real numbers. Prove that $|x + y| \leq |x| + |y|$.

Solution. First, every real number a satisfies $a \leq |a|$ and $-a \leq |a|$. This is itself a proof by cases [$\vee$E]: if $a \geq 0$, then $|a| = a$, so $a \leq |a|$, and $-a \leq 0 \leq a = |a|$; if $a < 0$, then $|a| = -a$, so $-a \leq |a|$, and $a < 0 < -a = |a|$.

Now $x + y \geq 0$ or $x + y < 0$, and we argue by cases [$\vee$E].

Case $x + y \geq 0$. Then $|x + y| = x + y \leq |x| + |y|$, adding $x \leq |x|$ and $y \leq |y|$.

Case $x + y < 0$. Then $|x + y| = -x - y = (-x) + (-y) \leq |x| + |y|$, adding $-x \leq |x|$ and $-y \leq |y|$.

Both cases reach $|x + y| \leq |x| + |y|$, so it holds.

Two case splits appear here, one after the other rather than one inside the other. The first proves a small lemma for an arbitrary number a, both halves at once, and its first half, $a \leq |a|$, is Example (A number is at most its absolute value) again. The second uses both halves of the lemma, for $a = x$ and for $a = y$. Splitting on the sign of $x + y$ instead of the signs of x and y is what saves two cases: the absolute value in the goal is $|x + y|$, and its sign is the only one the goal's left side cares about.

Warning. A case split must cover every possibility. A common faulty proof of the triangle inequality treats "$x \geq 0$ and $y \geq 0$" and "$x < 0$ and $y < 0$", checks both, and stops. The disjunction of these two cases is not a fact: $x = 1$, $y = -1$ satisfies neither. So there is no line $\varphi \vee \psi$ for $\vee$E to cite, and the rule was never legally applied. The missing mixed-sign cases are the interesting ones: at $x = 1$, $y = -1$ the inequality reads $0 \leq 2$, and it is strict.

Warning. Cases must **exhaust**, and a split that silently drops a value is a missing case in disguise. "Every real x satisfies $x > 0$ or $x < 0$" is false: $x = 0$ satisfies neither. A proof that splits into $x > 0$ and $x < 0$ cites a disjunction that is not available, and proves nothing about $x = 0$. The true disjunctions are "$x \geq 0$ or $x < 0$", and "$x > 0$, or $x = 0$, or $x < 0$".

Cases may overlap, however. "$x \geq 0$ or $x \leq 0$" is true for every real x, and $x = 0$ lies in both

cases. Nothing in ∨E forbids this: the rule asks that the goal be reached in each case, not that the cases be disjoint. (This is the inclusive "or" of Section 1.1.)

Warning. Because "or" is inclusive, knowing one side of a disjunction says nothing about the other. The form "from $P \vee Q$ and P, conclude $\neg Q$" is called **affirming a disjunct**, and it is bad. Let n be an integer, with P: "n is even", Q: "n is a multiple of 3"; at $n = 6$ both premises are true and $\neg Q$ is false. No step of ∨E does this: its cases end in one common statement, never in the falsity of a side.

Quick check. Each of the following begins a proof that $x^2 \geq 0$ for every real number x. Which case splits are exhaustive, so that ∨E may be applied?

(i) Case $x \geq 0$; case $x < 0$.

(ii) Case $x > 0$; case $x < 0$.

(iii) Case $x \geq 0$; case $x \leq 0$.

Solution.

(i) Exhaustive: every real number is ≥ 0 or < 0. (ii) Not exhaustive: $x = 0$ is in neither case, so the disjunction cited is false. The claim is still true at 0, but this split never proves it there. (iii) Exhaustive. The cases overlap at $x = 0$, which does no harm: $x^2 = 0 \geq 0$ is reached in both.

1.6.7 The rules of this chapter

First, as for each rule before it, the newest rule matched against statements larger than single letters.

Example. Instances of ∨E

Disjunction	First box	Second box	New line
$(P \wedge Q) \vee (P \wedge R)$	$P \wedge Q \ldots\ P \vee S$	$P \wedge R \ldots\ P \vee S$	$P \vee S$
$(P \to Q) \vee (R \wedge S)$	$P \to Q \ldots$ $(R \wedge S) \vee (P \to Q)$	$R \wedge S \ldots$ $(R \wedge S) \vee (P \to Q)$	$(R \wedge S) \vee (P \to Q)$
$(P \vee Q) \vee (Q \wedge S)$	$P \vee Q$ (a box of one line)	$Q \wedge S \ldots\ P \vee Q$	$P \vee Q$

Not an instance. From $(P \vee Q) \wedge R$, two boxes opened by P and by Q cannot be closed by ∨E citing that line: its main connective is ∧, so it is no disjunction. Take out $P \vee Q$ by ∧E first, and split on that line.

This chapter has introduced six rules, one or two at a time, and the structural clause R. Here is the whole chapter on one page.

Rule	Shape	You write	Lean
∧I	$\varphi, \psi \Longrightarrow \varphi \wedge \psi$	"We show P, and we show Q."	`constructor`
∧E	$\varphi \wedge \psi \Longrightarrow \varphi$ (or ψ)	"In particular, P."	`h.1` (and `h.2`)
→E	$\varphi \to \psi, \varphi \Longrightarrow \psi$	"Since P implies Q, and P holds, Q."	`exact h hp`
→I	box $\varphi \ldots \psi \Longrightarrow \varphi \to \psi$	"Assume P. … Hence Q."	`intro hp`
R	φ (available) $\Longrightarrow \varphi$, copied	"As shown above, P."	`exact hp`
∨I	$\varphi \Longrightarrow \varphi \vee \psi$ (or $\psi \vee \varphi$)	"So P or Q holds."	`left` (or `Or.inl hp`); `right`
∨E	$\varphi \vee \psi$, box $\varphi \ldots \chi$, box $\psi \ldots \chi \Longrightarrow \chi$	"Case 1: P. … so R. Case 2: Q. … so R. In either case, R."	`Or.elim h` (or `cases h`)

A written proof often drops the first part of the phrase for ∨E, saying "we argue by cases" without naming the disjunction; when that is not a standard fact such as "$x \geq 0$ or $x < 0$", ask for it, since it is the line the rule cites first. Nothing depends on the Lean column.

Warning (Not allowed: ∨E). Each step breaks the clause of the rule named after it.

- "Close the split with one case box": the clause "Two boxes".
- "Close it when the boxes end in R and in S": the clause "The same last line".
- "In the second box, cite a line of the first": the clause "Two boxes" and the scope rule; the boxes are separate (the warning after the rule).
- "Split on 'case $x > 0$, case $x < 0$' ": the clause "The disjunction"; the cases must be the sides of an available disjunction, and $(x > 0) \vee (x < 0)$ is not one (the warning on exhaustive cases).
- "Cite the boxes' lines one by one": the clause "Three citations"; each box is cited as a range.

Here is the third step as a derivation, wrong on the left and right on the right, with premises $(P \wedge Q) \vee R$ and $R \to Q$ and goal Q.

1	$(P \wedge Q) \vee R$	
2	$R \to Q$	
3	$\quad P \wedge Q$	asm, for ∨E
4	$\quad Q$	∧E 3
5	$\quad R$	asm, for ∨E
6	$\quad Q$	∧E 3
7	Q	∨E 1, 3–4, 5–6

1	$(P \wedge Q) \vee R$	
2	$R \to Q$	
3	$\quad P \wedge Q$	asm, for ∨E
4	$\quad Q$	∧E 3
5	$\quad R$	asm, for ∨E
6	$\quad Q$	→E 2, 5
7	Q	∨E 1, 3–4, 5–6

The wrong line 6 uses the first case's assumption in the second case. Its lines never cite premise 2, so they would "derive" Q from $(P \wedge Q) \vee R$ alone, and that claim breaks: with P: "$x > 0$", Q: "$x > 1$", R: "$x < 0$", at $x = -1$ the premise is true and Q is false.

Here are the rules as schemas. Letters in the number column stand for line numbers; every cited line must be available (Definition 1.4.2), and each box rule needs the box's last line as a line of that box itself.

Conjunction (Definition 1.2.5): ∧I, and the two forms of ∧E.

i	φ	
j	ψ	
k	$\varphi \wedge \psi$	$\wedge$I i, j

i	$\varphi \wedge \psi$	
k	φ	$\wedge$E i

i	$\varphi \wedge \psi$	
k	ψ	$\wedge$E i

Implication (Definition 1.4.1): →E and →I; and **repetition**, clause (D3) of Definition 1.4.2.

i	$\varphi \to \psi$	
j	φ	
k	ψ	→E i, j

i	φ	asm, for →I
	⋮	
j	ψ	
k	$\varphi \to \psi$	→I i–j

i	φ	
	⋮	
k	φ	R i

Disjunction (Definition 1.6.1): the two forms of ∨I, and ∨E.

i	φ	
k	$\varphi \vee \psi$	∨I i

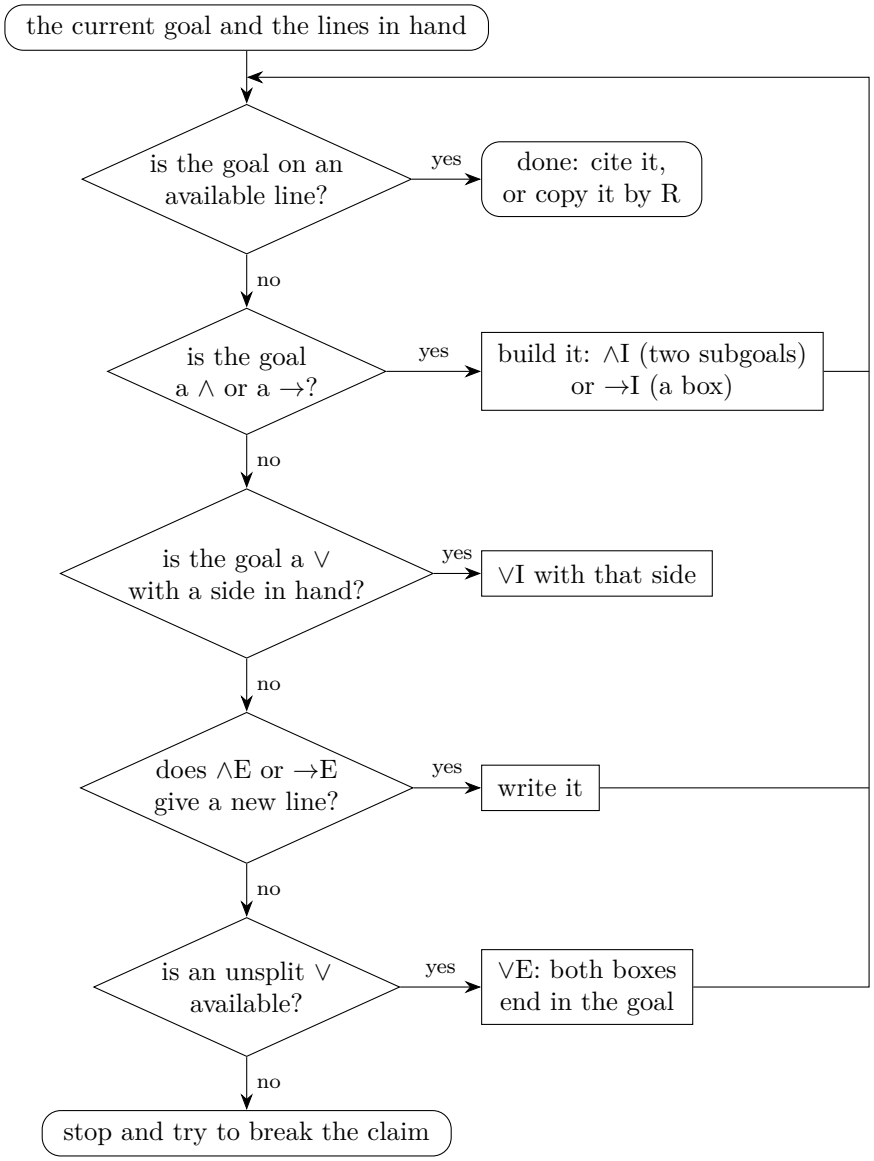
i	φ	
k	$\psi \vee \varphi$	∨I i

i	$\varphi \vee \psi$	
j	φ	asm, for ∨E
	⋮	
k	χ	
l	ψ	asm, for ∨E
	⋮	
m	χ	
n	χ	∨E i, j–k, l–m

Each rule is also a move in the search for a derivation. Read by the goal's main connective, the introduction rules say how a goal is built: that is the move *read the skeleton off the goal.* Read by an available line's main connective, the elimination rules say how that line is used: *use a premise by its main connective.*

Main connective	As the goal (goal-directed)	As an available line (premise-directed)
∧	two subgoals, one per side; finish with ∧I	take it apart with ∧E
→	open a box with the hypothesis; the new goal is the conclusion; close with →I	fire it with →E once its hypothesis is on a line, or make the hypothesis a subgoal
∨	postpone; choose a side with ∨I when that side is in hand	split with ∨E; both boxes end in the current goal

The table of moves as an order in which to try them:



The chart uses only this chapter's rules, and following it never writes an illegal line. It is a guide, not a rule: after asking whether the goal is already in hand, it reads the table's goal column and then its premise column, each from the top, except that it leaves out two moves. When an available implication ends in the goal but its hypothesis is on no line, make the hypothesis the current goal and start the chart again; items 3, 8 and 10 of Exercise (B7: Fifteen standard sequents) need this. When the goal is a disjunction and neither side is in hand, make one side the current goal; item 14 of Exercise (B7: Fifteen standard sequents) does this in each case. ("Unsplit" at the test for $\vee$ means not already split by a case box you are inside; without that word the chart would split $P \vee Q$ again in each case of item 14 and never stop.) So when the chart stops, try these moves. A claim the chart cannot finish may still be derivable: stopping is no evidence that the claim breaks.

1.6.7.1 Named moves

- **Try to break it.** To test a form, put statements in for the letters that make every premise true and the conclusion false; one such choice shows the form is bad. (Section 1.1)
- **Two ways to attack an argument.** Show a premise false, or show a step's form bad; the two attacks are independent. (Section 1.1)
- **Read the skeleton off the goal.** The goal's main connective chooses the introduction rule: $\wedge$ gives two subgoals, $\rightarrow$ a box, $\vee$ a choice to postpone. (Section 1.2)

- **Use a premise by its main connective.** Its elimination rule says what you may do with it: take a conjunction apart, feed an implication its hypothesis, split a disjunction into cases. (Section 1.2)
- **Open a box, close a box.** An assumption lives exactly as long as its box. (Section 1.4)
- **Work from both ends.** Premises at the top, the goal at the bottom; fill the gap from both sides. (Section 1.2)
- **Postpone a disjunction.** Choose a side of a goal $P \vee Q$ only when that side is in hand. (Section 1.5)
- **Cases end in the current goal.** Both boxes of $\vee$E close on the statement you are trying to reach. (Section 1.6)

Remark (Looking ahead). The rules for "not", for $\bot$, and for "if and only if", and the rule for proof by contradiction, are the subject of Chapter 2; the method that turns the table of moves above into a way of finding derivations is the subject of Chapter 3.

1.6.8 Exercises

1.6.8.1 A. Check your understanding

Exercise. A1

(a) State the rule $\vee$E as a schema, and name the three things its justification cites.

(b) For each sentence, taken from a proof, name the disjunction that supplies the cases and the statement both cases end in.
- "If n is even, then $n^2 + n = n(n+1)$ is even; if n is odd, then $n+1$ is even and so is $n(n+1)$. Either way, $n^2 + n$ is even."
- "Both possibilities for the sign of t give $t^2 \geq 0$."

(c) Let n be an integer. True or false, with a reason: "If n is even, then n^2 is even. Since every integer is even or odd, n^2 is even" is a correct proof by cases that n^2 is even.

(d) The following is offered as a derivation of $P \vee Q,\ P \to R \vdash R$. Name the line that is not justified and the clause of Definition 1.6.1 it violates.

1	$P \vee Q$	
2	$P \to R$	
3	$\quad P$	asm, for $\vee$E
4	$\quad R$	$\to$E 2, 3
5	R	$\vee$E 1, 3–4

(e) In the first derivation of Proposition 1.6.1, which line is the disjunction that $\vee$E cites at line 10, and where did it come from? Why does each box end in $(P \wedge Q) \vee (P \wedge R)$ rather than in $P \wedge Q$ and $P \wedge R$?

(f) In the same derivation, why does line 6 use $\vee$I to put $P \wedge Q$ on the **left**? Could line 6 have been $(P \wedge R) \vee (P \wedge Q)$?

Solution.

(a) The schema is the first one of Definition 1.6.1: from a line i holding $\varphi \vee \psi$, a box j–k

opened by φ, and a box l–m opened by ψ, each ending in the same statement χ as a line of its own, a line n holding χ. The justification "$\vee$E i, j–k, l–m" cites the disjunction and the two boxes.

(b) In the first, the disjunction is "n is even or n is odd", and both cases end in "$n^2 + n$ is even". In the second, it is "$t \geq 0$ or $t < 0$" (the sentence does not say it, and a careful reader supplies it), and both cases end in $t^2 \geq 0$.

(c) False. The disjunction "n is even or n is odd" does not say which side holds, so $\vee$E needs a second case, reaching "n^2 is even" from "n is odd", and the argument has none. Here the second case cannot be supplied: at $n = 1$ the number $n^2 = 1$ is odd, so the claim itself is false.

(d) Line 5. $\vee$E needs two boxes, one opened by each side of the disjunction (the clause "two boxes"); line 5 cites only the box opened by P, and no box opened by Q exists. Nor can a box opened by Q simply be added: nothing available reaches R from Q. In fact the sequent's form is bad, so no correct argument should be expected to reach R from these premises. Try to break it: for a real number x, take P: "$x > 1$", Q: "$x < 1$", R: "$x > 0$". At $x = -1$ both premises hold and R does not.

(e) Line 3, $Q \vee R$, obtained from the premise by $\wedge$E. Each box must end in the current goal, $(P \wedge Q) \vee (P \wedge R)$, because line 10 writes one statement, the same one that ends both boxes; boxes ending in $P \wedge Q$ and $P \wedge R$ would end in different statements, and $\vee$E could not be applied to them.

(f) Line 6 must be the current goal exactly, since both boxes end in it and line 10 writes it. The goal $(P \wedge Q) \vee (P \wedge R)$ has $P \wedge Q$ as its **left** side, so line 6 uses the first schema of $\vee$I. The line $(P \wedge R) \vee (P \wedge Q)$ is also an instance of $\vee$I on line 5 (the second schema), but it is a different statement: the first box would then end differently from the second, whose last line 9 is $(P \wedge Q) \vee (P \wedge R)$, and $\vee$E could not cite both boxes.

Exercise. A2: Drill

Each part takes a line or two. The parts run from reading a split, through a missing line and a justification column, to whole derivations. Let x be a real number and n an integer.

(a) "If $x \geq 0$, then $|x|^2 = x^2$; if $x < 0$, then $|x|^2 = (-x)^2 = x^2$. So $|x|^2 = x^2$." Name the disjunction that supplies the cases, and the common last line.

(b) How many case boxes does the proof of Example (The absolute value of a product) open?

(c) Inside a box whose current goal is S, the available line $Q \vee R$ is split. What do the two case boxes end in?

(d) Line 1 is $(P \wedge Q) \vee R$; a box 2–4 runs from $P \wedge Q$ to $S \vee P$, and a box 5–7 from R to $S \vee P$. What does "$\vee$E 1, 2–4, 5–7" write?

(e) Line 9 is Q, justified "$\vee$E 2, 3–5, 6–8", with line 3 $P \to Q$ and line 6 R. What are lines 2, 5 and 8?

(f) Justify each line: 1 $P \vee (Q \wedge R)$, 2 $P \to S$, 3 $R \to S$, premises; 4 P; 5 S; 6 $Q \wedge R$; 7 R; 8 S; 9 S. The boxes are 4–5 and 6–8.

(g) Two case boxes end in $R \wedge S$ and $S \wedge R$. May $\vee$E close the split?

(h) Find the illegal line.

1	$P \lor Q$	
2	$P \to R$	
3	$Q \to R$	
4	$\quad P$	asm, for $\lor$E
5	$\quad R$	$\to$E 2, 4
6	$\quad Q$	asm, for $\lor$E
7	$\quad R$	$\to$E 3, 4
8	R	$\lor$E 1, 4–5, 6–7

(i) Is the split "$n > 0$ or $n < 0$" exhaustive for integers?

(j) "Case $x \geq 1$: $x^2 \geq 1$. Case $x \leq -1$: $x^2 \geq 1$. So $x^2 \geq 1$ for every real x." What fails?

(k) Line 1 is $(P \lor Q) \to R$. A student opens a box with P and a box with Q, each ending in R, and writes "R — $\lor$E 1, 2–4, 5–7". Correct?

(l) Derive $P \lor Q \vdash Q \lor (P \lor R)$.

(m) Derive $(P \land Q) \lor (P \land R) \vdash P$.

(n) Derive $P \lor Q,\ Q \to P \vdash P$.

(o) Derive $(P \land Q) \lor (R \land S) \vdash Q \lor R$.

Solution.

(a) The disjunction is $(x \geq 0) \lor (x < 0)$, and both cases end in $|x|^2 = x^2$.

(b) Six: two for the sign of x, and two inside each of those for the sign of y. (The opening remark that $|a| = -a$ for $a \leq 0$ is itself a small split on $a < 0$ or $a = 0$, like the lemma in Example (The triangle inequality by cases); counting its two boxes as well gives eight.)

(c) S, the current goal at the point of the split.

(d) $S \lor P$, the common last line of the two boxes.

(e) Line 2 is $(P \to Q) \lor R$, the disjunction whose sides open the boxes at lines 3 and 6; lines 5 and 8 are both Q, the last line of each box.

(f) 4 assumption, for $\lor$E; 5 $\to$E 2, 4; 6 assumption, for $\lor$E; 7 $\land$E 6; 8 $\to$E 3, 7; 9 $\lor$E 1, 4–5, 6–8.

(g) No: both boxes must end in the same statement. Add $\land$E and $\land$I lines so that both end in the current goal, say $R \land S$.

(h) Line 7. It cites line 4, which lies in the box 4–5, closed before line 7. It should read "$\to$E 3, 6".

(i) No: $n = 0$ is in neither case.

(j) The split is not exhaustive: $(x \geq 1) \lor (x \leq -1)$ is false at $x = 0$, so $\lor$E has no disjunction to cite. The conclusion fails there too: $0^2 = 0 < 1$.

(k) No: the main connective of line 1 is $\to$, so it is no disjunction, and $\lor$E has nothing to split.

(l) Box 2–4: 2 P, assumption, for $\lor$E; 3 $P \lor R$, $\lor$I 2; 4 $Q \lor (P \lor R)$, $\lor$I 3. Box 5–6: 5 Q, assumption, for $\lor$E; 6 $Q \lor (P \lor R)$, $\lor$I 5. Then 7 $Q \lor (P \lor R)$, $\lor$E 1, 2–4, 5–6.

(m) Box 2–3: 2 $P \wedge Q$, assumption, for $\vee$E; 3 P, $\wedge$E 2. Box 4–5: 4 $P \wedge R$, assumption, for $\vee$E; 5 P, $\wedge$E 4. Then 6 P, $\vee$E 1, 2–3, 4–5.

(n) Box 3–3: 3 P, assumption, for $\vee$E, already the goal. Box 4–5: 4 Q, assumption, for $\vee$E; 5 P, $\to$E 2, 4. Then 6 P, $\vee$E 1, 3–3, 4–5.

(o) Box 2–4: 2 $P \wedge Q$, assumption, for $\vee$E; 3 Q, $\wedge$E 2; 4 $Q \vee R$, $\vee$I 3. Box 5–7: 5 $R \wedge S$, assumption, for $\vee$E; 6 R, $\wedge$E 5; 7 $Q \vee R$, $\vee$I 6. Then 8 $Q \vee R$, $\vee$E 1, 2–4, 5–7.

1.6.8.2 B. Practice

Exercise. B1: The other distributive law

Show that $(P \vee Q) \wedge (P \vee R) \vdash P \vee (Q \wedge R)$.

Hint: after taking the premise apart, you will need a case split inside a case.

Solution. The goal is a disjunction with no side in hand, so we postpone it. The premise gives two disjunctions. Split on $P \vee Q$, both cases ending in the goal. In the case P, the goal is one weakening away. In the case Q, we still lack R, so we split on $P \vee R$ inside that box: its case P weakens as before, and its case R combines with Q.

1	$(P \vee Q) \wedge (P \vee R)$	
2	$P \vee Q$	$\wedge$E 1
3	$P \vee R$	$\wedge$E 1
4	$\quad P$	asm, for $\vee$E
5	$\quad P \vee (Q \wedge R)$	$\vee$I 4
6	$\quad Q$	asm, for $\vee$E
7	$\quad\quad P$	asm, for $\vee$E
8	$\quad\quad P \vee (Q \wedge R)$	$\vee$I 7
9	$\quad\quad R$	asm, for $\vee$E
10	$\quad\quad Q \wedge R$	$\wedge$I 6, 9
11	$\quad\quad P \vee (Q \wedge R)$	$\vee$I 10
12	$\quad P \vee (Q \wedge R)$	$\vee$E 3, 7–8, 9–11
13	$P \vee (Q \wedge R)$	$\vee$E 2, 4–5, 6–12

Line 10 cites line 6, the assumption of the enclosing box, which is still open at line 10. Line 12 cites line 3, which lies in no box. Splitting on $P \vee R$ first, with the roles of Q and R exchanged, works equally well.

Exercise. B2: The larger of two numbers

For real numbers a and b, let $\max(a, b) := a$ if $a \geq b$, and $\max(a, b) := b$ if $a < b$. Prove that $\max(a, b) \geq a$ and $\max(a, b) \geq b$. Hence deduce that $\max(a, b) \geq \frac{1}{2}(a + b)$: the larger of two

numbers is at least their average.

Solution. We have $a \geq b$ or $a < b$, and we argue by cases [$\vee$E]; the current goal is the conjunction "$\max(a,b) \geq a$ and $\max(a,b) \geq b$".

Case $a \geq b$. Then $\max(a,b) = a$. So $\max(a,b) \geq a$, since $a = a$ and $\geq$ allows equality [$\vee$I]; and $\max(a,b) = a \geq b$. So both hold [$\wedge$I].

Case $a < b$. Then $\max(a,b) = b$. So $\max(a,b) = b > a$, which gives $\max(a,b) \geq a$ [$\vee$I]; and $\max(a,b) = b \geq b$ [$\vee$I]. So both hold [$\wedge$I].

In either case $\max(a,b) \geq a$ and $\max(a,b) \geq b$, so both hold.

Adding the two inequalities gives $2\max(a,b) \geq a+b$, and dividing by $2 > 0$ gives $\max(a,b) \geq \frac{1}{2}(a+b)$. For instance, $\max(3,-5) = 3 \geq -1 = \frac{1}{2}(3+(-5))$.

Exercise. B3: Two cases, two conclusions

(a) A search for $P \vee Q$, $P \to R$, $Q \to (S \wedge T) \vdash R \vee T$ has reached the stage below. *Name the move, then complete it.*

1	$P \vee Q$	
2	$P \to R$	
3	$Q \to (S \wedge T)$	
4	$\quad P$	asm, for $\vee$E
	$\quad \vdots$	
6	$\quad R \vee T$	
7	$\quad Q$	asm, for $\vee$E
	$\quad \vdots$	
10	$\quad R \vee T$	
11	$R \vee T$	?

(b) Use (a) to prove: for every integer n, n^2 is a multiple of 4 or one more than a multiple of 8. You may use that every integer is even or odd, and that $k(k+1)$ is even for every integer k.

Solution.

(a) The move is to **split on the premise** $P \vee Q$, each case ending in the current goal: *cases end in the current goal.* The goal $R \vee T$ is a disjunction with no side in hand, so it is postponed, and line 11 will be $\vee$E citing line 1 and the two case boxes. In the case P, line 2 fires and $\vee$I weakens R to the goal. In the case Q, line 3 fires, $\wedge$E takes out T, and $\vee$I weakens it to the goal on the other side.

1	$P \vee Q$	
2	$P \to R$	
3	$Q \to (S \wedge T)$	
4	$\quad P$	asm, for $\vee$E
5	$\quad R$	$\to$E 2, 4
6	$\quad R \vee T$	$\vee$I 5
7	$\quad Q$	asm, for $\vee$E
8	$\quad S \wedge T$	$\to$E 3, 7
9	$\quad T$	$\wedge$E 8
10	$\quad R \vee T$	$\vee$I 9
11	$R \vee T$	$\vee$E 1, 4–6, 7–10

The two boxes reach different statements, R at line 5 and T at line 9, and $\vee$I turns both into the common last line $R \vee T$.

(b) Take P: "n is even", Q: "n is odd", R: "n^2 is a multiple of 4", S: "n^2 is odd", T: "n^2 is one more than a multiple of 8". The premise $P \vee Q$ is the fact we may use. For $P \to R$: if $n = 2k$ for an integer k, then $n^2 = 4k^2$. For $Q \to (S \wedge T)$: if $n = 2k+1$, then $n^2 = 4k^2 + 4k + 1 = 4k(k+1) + 1$, which is odd; and with $k(k+1) = 2m$ for an integer m, $n^2 = 8m + 1$. By (a), $R \vee T$. For instance, $6^2 = 36 = 4 \cdot 9$, and $7^2 = 49 = 8 \cdot 6 + 1$.

Exercise. B4: The smaller of two numbers

For real numbers a and b, let $\max(a, b)$ be as in Exercise (B2: The larger of two numbers), and let $\min(a, b) := b$ if $a \geq b$, and $\min(a, b) := a$ if $a < b$. Prove by cases that $\min(a, b) = -\max(-a, -b)$.

Solution. We have $a \geq b$ or $a < b$, and we argue by cases [$\vee$E]; the current goal is the equation $\min(a, b) = -\max(-a, -b)$.

Case $a < b$. Then $\min(a, b) = a$. Also $-a > -b$, so $-a \geq -b$ [$\vee$I], and by the definition of max, $\max(-a, -b) = -a$. Its negative is $a = \min(a, b)$.

Case $a \geq b$. Then $\min(a, b) = b$, and $-a \leq -b$. Written out, as the convention on inequalities of Section 1.5 asks when a case split must see the disjunction, this is $(-a < -b) \vee (-a = -b)$, and we split again [$\vee$E].

- If $-a < -b$, then $\max(-a, -b) = -b$, whose negative is $b = \min(a, b)$.
- If $-a = -b$, then $-a \geq -b$ [$\vee$I], so $\max(-a, -b) = -a$, whose negative is a; and $a = b = \min(a, b)$.

Both inner cases reach the goal, so $\vee$E closes the inner split; then both outer cases have reached it, and $\vee$E closes the outer split. Hence $\min(a, b) = -\max(-a, -b)$.

For instance, at $a = 3$, $b = -5$: $\min(3, -5) = -5$, and $\max(-3, 5) = 5$, whose negative is -5. At $a = b = 2$: $\min(2, 2) = 2$, and $\max(-2, -2) = -2$, whose negative is 2.

Exercise. B5: Derive or break

For each claim, derive it with the rules of this chapter, or break it by statements about a real number x that make its counterexample set true.

(a) $P \vee Q,\ R \vdash (P \wedge R) \vee Q$.
(b) $(P \vee Q) \wedge (P \vee R) \vdash P \wedge (Q \vee R)$.
(c) $P \vee (Q \vee R),\ Q \to P \vdash P \vee R$.
(d) $(P \vee Q) \to R \vdash (P \to R) \vee (Q \to R)$.
(e) $(P \to R) \vee (Q \to R) \vdash (P \vee Q) \to R$.
(f) $P \vee Q,\ P \to R \vdash R \vee Q$.
(g) $P \vee (Q \to R),\ Q \vdash P \vee R$.
(h) $P \vee (Q \to R) \vdash (P \vee Q) \to (P \vee R)$.
(i) $(P \vee Q) \to (P \vee R) \vdash Q \to R$.

Solution.

(a) Derive. Split on line 1; in the case P, build $P \wedge R$ with the premise R.

1	$P \vee Q$	
2	R	
3	$\quad P$	asm, for $\vee$E
4	$\quad P \wedge R$	$\wedge$I 3, 2
5	$\quad (P \wedge R) \vee Q$	$\vee$I 4
6	$\quad Q$	asm, for $\vee$E
7	$\quad (P \wedge R) \vee Q$	$\vee$I 6
8	$(P \wedge R) \vee Q$	$\vee$E 1, 3–5, 6–7

(b) Break. Take P: "$x > 5$", Q: "$x > 0$", R: "$x < 10$". At $x = 1$, $P \vee Q$ and $P \vee R$ are true ($1 > 0$, $1 < 10$) and P is false, so $P \wedge (Q \vee R)$ is false: the counterexample set $(P \vee Q) \wedge (P \vee R)$, $\neg(P \wedge (Q \vee R))$ is true.

(c) Derive, with a split inside a case: the case $Q \vee R$ is itself a disjunction.

1	$P \vee (Q \vee R)$	
2	$Q \to P$	
3	$\quad P$	asm, for $\vee$E
4	$\quad P \vee R$	$\vee$I 3
5	$\quad Q \vee R$	asm, for $\vee$E
6	$\quad\quad Q$	asm, for $\vee$E
7	$\quad\quad P$	$\to$E 2, 6
8	$\quad\quad P \vee R$	$\vee$I 7
9	$\quad\quad R$	asm, for $\vee$E
10	$\quad\quad P \vee R$	$\vee$I 9
11	$\quad P \vee R$	$\vee$E 5, 6–8, 9–10
12	$P \vee R$	$\vee$E 1, 3–4, 5–11

(d) Derive. The goal is a disjunction; the left side $P \to R$ can be built, as in Example (A disjunctive hypothesis splits into two implications), and $\vee$I then adds the other side.

1	$(P \vee Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad R$	$\to$E 1, 3
5	$P \to R$	$\to$I 2–4
6	$(P \to R) \vee (Q \to R)$	$\vee$I 5

(e) Break. Take P: "$x < 0$", Q: "$x > 0$", R: "$x > 5$". At $x = 1$, $P \to R$ is true, since its hypothesis fails, so the premise is true; $P \vee Q$ is true and R false, so $(P \vee Q) \to R$ is false. The counterexample set is true.

(f) Derive.

1	$P \vee Q$	
2	$P \to R$	
3	$\quad P$	asm, for $\vee$E
4	$\quad R$	$\to$E 2, 3
5	$\quad R \vee Q$	$\vee$I 4
6	$\quad Q$	asm, for $\vee$E
7	$\quad R \vee Q$	$\vee$I 6
8	$R \vee Q$	$\vee$E 1, 3–5, 6–7

(g) Derive. In the second case the case's assumption is an implication, and the premise Q fires it.

1	$P \vee (Q \to R)$	
2	$\quad Q$	
3	$\quad\quad P$	asm, for $\vee$E
4	$\quad\quad P \vee R$	$\vee$I 3
5	$\quad\quad Q \to R$	asm, for $\vee$E
6	$\quad\quad R$	$\to$E 5, 2
7	$\quad\quad P \vee R$	$\vee$I 6
8	$P \vee R$	$\vee$E 1, 3–4, 5–7

(h) Derive, with a split inside a case, inside a box. The goal is an implication: open the box of $P \vee Q$. Inside it, split on the premise; in its case $Q \to R$, split again on the box's assumption $P \vee Q$, which is available there.

1	$P \vee (Q \to R)$	
2	$\quad P \vee Q$	asm, for $\to$I
3	$\quad\quad P$	asm, for $\vee$E
4	$\quad\quad P \vee R$	$\vee$I 3
5	$\quad\quad Q \to R$	asm, for $\vee$E
6	$\quad\quad\quad P$	asm, for $\vee$E
7	$\quad\quad\quad P \vee R$	$\vee$I 6
8	$\quad\quad\quad Q$	asm, for $\vee$E
9	$\quad\quad\quad R$	$\to$E 5, 8
10	$\quad\quad\quad P \vee R$	$\vee$I 9
11	$\quad\quad P \vee R$	$\vee$E 2, 6–7, 8–10
12	$\quad P \vee R$	$\vee$E 1, 3–4, 5–11
13	$(P \vee Q) \to (P \vee R)$	$\to$I 2–12

Line 11 cites line 2, the assumption of the box 2–12, which contains line 11; the boxes 3–4 and 5–11 sit directly in that box, as line 12 does.

(i) Break. Take P: "$x > 0$", Q: "$x > 1$", R: "$x > 5$". At $x = 2$, $P \vee R$ is true, so the premise $(P \vee Q) \to (P \vee R)$ is true; Q is true and R false, so $Q \to R$ is false. The counterexample set is true.

(ii) and (e) are converses: the original derives and the converse breaks. (i) is a look-alike that drops the shared P from (h)'s conclusion, and breaks. (b) is a look-alike of Exercise (B1: The other distributive law), with the conclusion's $\wedge$ and $\vee$ traded.

Exercise. B6: Translate, then derive

Translate each argument into a sequent with the letters given, then derive it or break its

form.

(a) Let n be an integer. *n is even or n is odd. If n is even, then $n(n+1)$ is even. If n is odd, then $n+1$ is even. If $n+1$ is even, then $n(n+1)$ is even. So $n(n+1)$ is even.* Letters: E: "n is even", O: "n is odd", F: "$n+1$ is even", M: "$n(n+1)$ is even".

(b) Let x be a real number. *If $x > 1$, then $x^2 > 1$. If $x < -1$, then $x^2 > 1$. And $x^2 > 1$. So $x > 1$ or $x < -1$.* Letters: P: "$x > 1$", Q: "$x < -1$", R: "$x^2 > 1$". If the form breaks, say whether the conclusion is still true for every x with $x^2 > 1$, and which premise would repair the argument.

Solution.

(a) The sequent is $E \lor O,\ E \to M,\ O \to F,\ F \to M \vdash M$. Split on the first premise; in the case O the chain of two implications reaches M.

1	$E \lor O$	
2	$E \to M$	
3	$O \to F$	
4	$F \to M$	
5	$\quad E$	asm, for $\lor$E
6	$\quad M$	$\to$E 2, 5
7	$\quad O$	asm, for $\lor$E
8	$\quad F$	$\to$E 3, 7
9	$\quad M$	$\to$E 4, 8
10	M	$\lor$E 1, 5–6, 7–9

(b) The sequent is $P \to R,\ Q \to R,\ R \vdash P \lor Q$, and its form breaks: it affirms the consequent of two implications at once. Keep P and Q, and replace R by "$x \neq 0$". Both implications then hold for every x, and at $x = \frac{1}{2}$ the premise R is true and $P \lor Q$ is false, so the counterexample set $P \to R,\ Q \to R,\ R,\ \neg(P \lor Q)$ is true.

With the argument's own content no number breaks it, because $x^2 > 1$ does imply $x > 1$ or $x < -1$ (Exercise (B3: Necessary and sufficient, used later), part (c)). Adding that fact as the premise $R \to (P \lor Q)$ gives a one-step derivation, $\to$E on it and the premise R, and makes the other two premises unnecessary. A true conclusion, a bad form: the argument proves nothing until the missing premise is written.

Exercise. B7: Fifteen standard sequents

Derive each sequent. They are grouped by the move that does the main work.

Read the skeleton off the goal.

1. $Q \vdash (P \to R) \to (P \to (Q \land R))$.
2. $\vdash (P \land Q) \to ((Q \to R) \to (R \land P))$.
3. $(P \lor Q) \to R \vdash P \to (S \to (R \land P))$.

4. $P \to (Q \to R) \vdash (S \wedge P) \to (Q \to (R \vee S))$.
5. $Q \to R \vdash P \to ((P \to Q) \to (R \wedge P))$.

Use a premise by its main connective.

6. $P \to (Q \wedge R) \vdash (P \wedge S) \to (R \wedge S)$.
7. $P \wedge (Q \to R),\ P \to Q \vdash R \wedge P$.
8. $(P \wedge Q) \to R,\ P \to Q \vdash P \to R$.
9. $P \to (Q \to R),\ S \to Q \vdash (P \wedge S) \to R$.
10. $((P \wedge Q) \to R) \to S,\ P \to T,\ (T \wedge Q) \to R \vdash S$.

Cases end in the current goal.

11. $(P \to Q) \vee (Q \wedge R),\ P \vdash Q$.
12. $(P \to R) \vee (Q \to R) \vdash (P \wedge Q) \to R$.
13. $(P \to Q) \vee (P \to R) \vdash P \to (Q \vee R)$.
14. $P \vee Q,\ S \to (Q \to R) \vdash (S \to (P \wedge S)) \vee (S \to R)$.
15. $P \vee (Q \wedge R),\ P \to (R \wedge S) \vdash R \wedge (S \vee Q)$.

Solution. **1.** Two arrows at the top of the goal give two nested boxes; inside, R comes by $\to$E and joins the premise.

1	Q	
2	$P \to R$	asm, for $\to$I
3	P	asm, for $\to$I
4	R	$\to$E 2, 3
5	$Q \wedge R$	$\wedge$I 1, 4
6	$P \to (Q \wedge R)$	$\to$I 3–5
7	$(P \to R) \to (P \to (Q \wedge R))$	$\to$I 2–6

2. No premises: the skeleton of the goal is two boxes, and everything inside comes from their assumptions.

1	$P \wedge Q$	asm, for $\to$I
2	$Q \to R$	asm, for $\to$I
3	Q	$\wedge$E 1
4	R	$\to$E 2, 3
5	P	$\wedge$E 1
6	$R \wedge P$	$\wedge$I 4, 5
7	$(Q \to R) \to (R \wedge P)$	$\to$I 2–6
8	$(P \wedge Q) \to ((Q \to R) \to (R \wedge P))$	$\to$I 1–7

3. Two boxes from the goal; inside, $\vee$I manufactures the premise's hypothesis. The assump-

tion S is never cited.

1	$(P \vee Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad S$	asm, for $\to$I
4	$\quad\quad P \vee Q$	$\vee$I 2
5	$\quad\quad R$	$\to$E 1, 4
6	$\quad\quad R \wedge P$	$\wedge$I 5, 2
7	$\quad S \to (R \wedge P)$	$\to$I 3–6
8	$P \to (S \to (R \wedge P))$	$\to$I 2–7

4. Two boxes from the goal; the premise fires on a half of the outer assumption, and the implication it gives fires on the inner one.

1	$P \to (Q \to R)$	
2	$\quad S \wedge P$	asm, for $\to$I
3	$\quad\quad Q$	asm, for $\to$I
4	$\quad\quad P$	$\wedge$E 2
5	$\quad\quad Q \to R$	$\to$E 1, 4
6	$\quad\quad R$	$\to$E 5, 3
7	$\quad\quad R \vee S$	$\vee$I 6
8	$\quad Q \to (R \vee S)$	$\to$I 3–7
9	$(S \wedge P) \to (Q \to (R \vee S))$	$\to$I 2–8

5. Two boxes from the goal; the inner assumption fires on the outer one, and the premise on the result.

1	$Q \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad P \to Q$	asm, for $\to$I
4	$\quad\quad Q$	$\to$E 3, 2
5	$\quad\quad R$	$\to$E 1, 4
6	$\quad\quad R \wedge P$	$\wedge$I 5, 2
7	$\quad (P \to Q) \to (R \wedge P)$	$\to$I 3–6
8	$P \to ((P \to Q) \to (R \wedge P))$	$\to$I 2–7

6. One box from the goal; inside, each line is read off a main connective: $\wedge$ is taken apart, $\to$ is fired.

1	$P \to (Q \wedge R)$	
2	$\quad P \wedge S$	asm, for $\to$I
3	$\quad P$	$\wedge$E 2
4	$\quad Q \wedge R$	$\to$E 1, 3
5	$\quad R$	$\wedge$E 4
6	$\quad S$	$\wedge$E 2
7	$\quad R \wedge S$	$\wedge$I 5, 6
8	$(P \wedge S) \to (R \wedge S)$	$\to$I 2–7

7. No box: the conjunction is taken apart, and its two halves feed the two implications.

1	$P \wedge (Q \to R)$	
2	$P \to Q$	
3	P	$\wedge$E 1
4	$Q \to R$	$\wedge$E 1
5	Q	$\to$E 2, 3
6	R	$\to$E 4, 5
7	$R \wedge P$	$\wedge$I 6, 3

8. The first premise needs its whole hypothesis $P \wedge Q$; the box gives P, and the second premise gives Q.

1	$(P \wedge Q) \to R$	
2	$P \to Q$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q$	$\to$E 2, 3
5	$\quad P \wedge Q$	$\wedge$I 3, 4
6	$\quad R$	$\to$E 1, 5
7	$P \to R$	$\to$I 3–6

9. Both halves of the assumption are used, one by each premise, and the results meet in a third $\to$E.

1	$P \to (Q \to R)$	
2	$S \to Q$	
3	$\quad P \wedge S$	asm, for $\to$I
4	$\quad P$	$\wedge$E 3
5	$\quad S$	$\wedge$E 3
6	$\quad Q \to R$	$\to$E 1, 4
7	$\quad Q$	$\to$E 2, 5
8	$\quad R$	$\to$E 6, 7
9	$(P \wedge S) \to R$	$\to$I 3–8

10. The only line mentioning S is the first premise, whose hypothesis is itself an implication: it becomes a subgoal, proved in a box by the other two premises.

1	$((P \wedge Q) \to R) \to S$	
2	$P \to T$	
3	$(T \wedge Q) \to R$	
4	$\quad P \wedge Q$	asm, for $\to$I
5	$\quad P$	$\wedge$E 4
6	$\quad T$	$\to$E 2, 5
7	$\quad Q$	$\wedge$E 4
8	$\quad T \wedge Q$	$\wedge$I 6, 7
9	$\quad R$	$\to$E 3, 8
10	$(P \wedge Q) \to R$	$\to$I 4–9
11	S	$\to$E 1, 10

11. Split on the first premise; in one case the premise P fires the case's implication, in the other $\wedge$E finishes.

1	$(P \to Q) \vee (Q \wedge R)$	
2	P	
3	$\quad P \to Q$	asm, for $\vee$E
4	$\quad Q$	$\to$E 3, 2
5	$\quad Q \wedge R$	asm, for $\vee$E
6	$\quad Q$	$\wedge$E 5
7	Q	$\vee$E 1, 3–4, 5–6

12. A box from the goal, then a split inside it, each case ending in the current goal R.

1	$(P \to R) \vee (Q \to R)$	
2	$\quad P \wedge Q$	asm, for $\to$I
3	$\quad\quad P \to R$	asm, for $\vee$E
4	$\quad\quad P$	$\wedge$E 2
5	$\quad\quad R$	$\to$E 3, 4
6	$\quad\quad Q \to R$	asm, for $\vee$E
7	$\quad\quad Q$	$\wedge$E 2
8	$\quad\quad R$	$\to$E 6, 7
9	$\quad R$	$\vee$E 1, 3–5, 6–8
10	$(P \wedge Q) \to R$	$\to$I 2–9

13. A box from the goal, then a split inside it; the current goal $Q \vee R$ is reached from a different side in each case.

1	$(P \to Q) \vee (P \to R)$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad P \to Q$	asm, for $\vee$E
4	$\quad\quad Q$	$\to$E 3, 2
5	$\quad\quad Q \vee R$	$\vee$I 4
6	$\quad\quad P \to R$	asm, for $\vee$E
7	$\quad\quad R$	$\to$E 6, 2
8	$\quad\quad Q \vee R$	$\vee$I 7
9	$\quad Q \vee R$	$\vee$E 1, 3–5, 6–8
10	$P \to (Q \vee R)$	$\to$I 2–9

14. The goal is a disjunction, so no box can be opened first: the split comes first, as in Exercise (B3: Two cases, two conclusions) (a), and each case opens its own $\to$I box. In the case Q, the box's assumption S is what opens the premise: $\to$E on line 2 needs S, which only the box supplies, and only then does the case's assumption Q give R.

1	$P \vee Q$	
2	$S \to (Q \to R)$	
3	$\quad P$	asm, for $\vee$E
4	$\quad\quad S$	asm, for $\to$I
5	$\quad\quad P \wedge S$	$\wedge$I 3, 4
6	$\quad S \to (P \wedge S)$	$\to$I 4–5
7	$\quad (S \to (P \wedge S)) \vee (S \to R)$	$\vee$I 6
8	$\quad Q$	asm, for $\vee$E
9	$\quad\quad S$	asm, for $\to$I
10	$\quad\quad Q \to R$	$\to$E 2, 9
11	$\quad\quad R$	$\to$E 10, 8
12	$\quad S \to R$	$\to$I 9–11
13	$\quad (S \to (P \wedge S)) \vee (S \to R)$	$\vee$I 12
14	$(S \to (P \wedge S)) \vee (S \to R)$	$\vee$E 1, 3–7, 8–13

15. Split on the first premise; each case builds both conjuncts of the goal, and $S \vee Q$ is reached from a different side in each.

1	$P \vee (Q \wedge R)$	
2	$P \to (R \wedge S)$	
3	$\quad P$	asm, for $\vee$E
4	$\quad R \wedge S$	$\to$E 2, 3
5	$\quad R$	$\wedge$E 4
6	$\quad S$	$\wedge$E 4
7	$\quad S \vee Q$	$\vee$I 6
8	$\quad R \wedge (S \vee Q)$	$\wedge$I 5, 7
9	$\quad Q \wedge R$	asm, for $\vee$E
10	$\quad Q$	$\wedge$E 9
11	$\quad R$	$\wedge$E 9
12	$\quad S \vee Q$	$\vee$I 10
13	$\quad R \wedge (S \vee Q)$	$\wedge$I 11, 12
14	$R \wedge (S \vee Q)$	$\vee$E 1, 3–8, 9–13

1.6.8.3 C. Going deeper

Exercise. C1: Challenge: The reverse triangle inequality

Let x and y be real numbers. Prove that

$$||x| - |y|| \leq |x - y|$$

by splitting into the two cases $|x| \geq |y|$ and $|x| < |y|$, using in each case the triangle inequality of Example (The triangle inequality by cases). Then explain why these two cases exhaust all possibilities, and why a split on the signs of x and y would need four cases.

Solution (Hint 1). The triangle inequality bounds $|u + v|$. Write x as a sum of two numbers whose absolute values appear in the goal.

Solution (Hint 2). In the case $|x| \geq |y|$, the left side is $|x| - |y|$. Start from $x = (x - y) + y$.

Solution. We use two school facts: for real numbers a, b, either $a \geq b$ or $a < b$; and $|-a| = |a|$ for every real a (if $a \geq 0$ then $-a \leq 0$, so $|-a| = -(-a) = a = |a|$, using, as in Example (The absolute value of a product), that $|b| = -b$ whenever $b \leq 0$; if $a < 0$ then $-a > 0$, so $|-a| = -a = |a|$).

By the first fact, $|x| \geq |y|$ or $|x| < |y|$, and we argue by cases [$\vee$E].

Case $|x| \geq |y|$. Then $|x| - |y| \geq 0$, so $||x| - |y|| = |x| - |y|$. By Example (The triangle inequality by cases) applied to the numbers $x - y$ and y,

$$|x| = |(x - y) + y| \leq |x - y| + |y|,$$

so $|x| - |y| \leq |x - y|$, which is the goal in this case.

Case $|x| < |y|$. Then $|x| - |y| < 0$, so $||x| - |y|| = |y| - |x|$. By Example (The triangle inequality by cases) applied to $y - x$ and x,

$$|y| = |(y - x) + x| \leq |y - x| + |x|,$$

so $|y| - |x| \leq |y - x| = |-(x - y)| = |x - y|$.

Both cases reach the goal, which proves the inequality.

The two cases exhaust all possibilities because they are the two sides of the disjunction "$a \geq b$ or $a < b$" with $a = |x|$, $b = |y|$, and that disjunction holds for all real numbers. The split was chosen by the goal: the outer absolute value is of $|x| - |y|$, and its formula depends only on the sign of that difference.

A split on signs would decide $|x|$ and $|y|$ separately: $x \geq 0$ or $x < 0$ gives the formula for $|x|$, and inside each case $y \geq 0$ or $y < 0$ gives the formula for $|y|$. That is a case split inside each box of another, with $2 \cdot 2 = 4$ inner boxes, as in Example (The absolute value of a product). Fewer would not do for that method: with only the sign of x fixed, $|y|$ still has no formula.

Marking scheme. (10 points)

- 2: states the disjunction $|x| \geq |y|$ or $|x| < |y|$ and says it holds for all reals.

- 3: in the first case, removes the outer absolute value correctly and derives $|x| - |y| \leq |x - y|$ from the triangle inequality for $(x - y) + y$.
- 3: in the second case, the same with the roles exchanged, including $|y - x| = |x - y|$.
- 2: explains the four sign cases as a split inside each case of another.

Fatal errors (0 for the step): using $||x| - |y|| = |x| - |y|$ in both cases; applying the triangle inequality to a difference as if $|u - v| \leq |u| - |v|$. A proof that squares both sides instead, or splits on signs, earns the 6 points of the two middle lines if every case is complete.

Exercise. C2: Find the error

Here is a proof, by four cases, that $|x - y| \leq |x| + |y|$ for all real numbers x and y.

> *Case 1:* $x \geq 0$. Then $|x| = x$.
>
> *Case 1a:* $y \geq 0$. Then $|y| = y$, and $|x-y|$ is $x-y$ or $y-x$. Since $x-y \leq x \leq x+y$ and $y - x \leq y \leq x + y$, we get $|x - y| \leq x + y = |x| + |y|$.
>
> *Case 1b:* $y < 0$. Then $|y| = -y$, and $x - y > 0$, so $|x - y| = x - y = |x| + |y|$.
>
> *Case 2:* $x < 0$. Then $|x| = -x$.
>
> *Case 2a:* $y \geq 0$. Then $|y| = y$, and $x - y < 0$, so $|x - y| = y - x = |y| + |x|$.
>
> *Case 2b:* $y < 0$. Then $|y| = -y$. Since $x \geq 0$ and $y < 0$, we have $x - y > 0$, so $|x - y| = x - y \leq |x| + |y|$, because $x \leq |x|$ and $-y = |y|$.
>
> All four cases give $|x - y| \leq |x| + |y|$.

(a) Is the statement true?

(b) Find the line that is not justified, and name the rule whose conditions it breaks.

(c) Show with specific numbers that the step fails.

(d) Repair Case 2b.

Solution.

(a) Yes. For instance, it follows from Example (The triangle inequality by cases) applied to x and $-y$, since $|-y| = |y|$ (shown in the solution of Exercise (C1: Challenge: The reverse triangle inequality)). So the question is not whether the conclusion holds but whether this argument proves it.

(b) The sentence "Since $x \geq 0$ and $y < 0$, we have $x - y > 0$" in Case 2b. The assumption $x \geq 0$ is the assumption of Case 1. In derivation terms, it opens the first box of the outer $\vee$E, and Case 2b lies inside the **second** box, which is opened by $x < 0$. Once the first box has closed, its lines are not available (the scope rule of Definition 1.4.2), so Case 2b may not cite $x \geq 0$: the rule broken is $\vee$E, each of whose boxes may cite only lines available at it by the scope rule: its own assumption, what has been derived inside it, and what was available before the box opened; so nothing from the first box, which has closed. The sentence is the reasoning of Case 1b copied into Case 2b, where its assumption is no longer in force. So the box of Case 2 does not reach the goal from the lines available to it, and the final $\vee$E has no right to cite it.

(c) Take $x = -3$ and $y = -1$, which is Case 2b. Then $x - y = -2$, which is not > 0, and

$|x-y| = 2 \neq -2 = x-y$. The conclusion of Case 2b still holds here ($2 \leq 3+1$), but the step used to reach it is false.

(d) In Case 2b, $x < 0$ and $y < 0$, so $|x| = -x$ and $|y| = -y$. The number $|x-y|$ is $x-y$ or $y-x$. Since $x < 0$, we have $x - y < -y = |y| \leq |x| + |y|$. Since $y < 0$, we have $y - x < -x = |x| \leq |x| + |y|$. So in either case $|x-y| \leq |x| + |y|$, using only the assumptions of Case 2b.

Exercise. C3: Exit check: a distributive law, rule by rule

Derive $(P \wedge Q) \vee (R \wedge Q) \vdash (P \vee R) \wedge Q$, naming the rule and the cited lines on every line. Then say which steps would change if the premise were $(P \wedge Q) \vee (P \wedge R)$ and the goal $P \wedge (Q \vee R)$, and which would stay the same.

Solution. The chart would build the conjunction by ∧I, with subgoals $P \vee R$ and Q, and reach each one by splitting the premise; that route works (it is the second route below). Here one split is enough: the premise is a disjunction, so we split on it first and build the whole goal in each case; both cases end in the current goal $(P \vee R) \wedge Q$. In each case, ∧E takes the assumption apart, ∨I builds $P \vee R$ from the side in hand, and ∧I builds the goal.

1	$(P \wedge Q) \vee (R \wedge Q)$	
2	$\quad P \wedge Q$	asm, for ∨E
3	$\quad P$	∧E 2
4	$\quad Q$	∧E 2
5	$\quad P \vee R$	∨I 3
6	$\quad (P \vee R) \wedge Q$	∧I 5, 4
7	$\quad R \wedge Q$	asm, for ∨E
8	$\quad R$	∧E 7
9	$\quad Q$	∧E 7
10	$\quad P \vee R$	∨I 8
11	$\quad (P \vee R) \wedge Q$	∧I 10, 9
12	$(P \vee R) \wedge Q$	∨E 1, 2–6, 7–11

Line by line: 1 is the premise; 2 is the assumption of the first case; 3 and 4 are ∧E on line 2; 5 is ∨I on line 3, first schema; 6 is ∧I on lines 5 and 4, written with $P \vee R$ as the left conjunct; 7 is the assumption of the second case; 8 and 9 are ∧E on line 7; 10 is ∨I on line 8, second schema; 11 is ∧I on lines 10 and 9; 12 is ∨E, citing the disjunction on line 1 and the boxes 2–6 and 7–11, both of which end in the goal as lines of their own. No line of the second box cites the first.

With premise $(P \wedge Q) \vee (P \wedge R)$ and goal $P \wedge (Q \vee R)$, the shape stays the same: a split on the premise, ∧E twice in each box, one ∨I and one ∧I in each box, and ∨E at the end. What changes is which part is weakened and where it goes. The common conjunct is now P, and the parts that differ between the cases are Q and R; so the ∨I lines build $Q \vee R$ (from Q in

the first box, first schema; from R in the second, second schema), and the $\wedge$I lines write P as the left conjunct, as in the goal. That derivation is the second one of Proposition 1.6.1.

Marking scheme. (10 points)

- 2: splits on the premise by $\vee$E, with two boxes opened by its two sides.
- 2: in each box, takes the assumption apart by $\wedge$E into both conjuncts.
- 2: in each box, uses $\vee$I to build $P \vee R$ from the side in hand, with the correct schema.
- 2: in each box, builds the goal by $\wedge$I, so that both boxes end in the same line, the goal.
- 2: the $\vee$E line cites the disjunction and both boxes as ranges, and the final sentences name the changed steps (the $\vee$I lines, now building $Q \vee R$, and the $\wedge$I lines, now with P on the left).

Fatal errors (0 for the step): a line in one box citing a line of the other box; a conclusion drawn inside one box only and then used outside it.

A second route is also correct: split the premise twice, once ending both boxes in $P \vee R$ and once ending both in Q, and then join the two results by $\wedge$I outside every box (14 lines). Its points map as follows: 2 for either split with two boxes opened by the premise's sides; 2 for the $\wedge$E lines in the boxes; 2 for the $\vee$I sides; 2 for the final $\wedge$I, with the boxes of each split ending in the same line; 2 for the citations and the final sentences.

Exercise. C4: The second half

A search for $Q \vee R,\ Q \to (P \to S),\ R \to S \vdash P \to S$ has written the premises and opened one box; the rest is blank. The finished derivation has 11 lines. *Name the move, then complete it.*

1	$Q \vee R$	
2	$Q \to (P \to S)$	
3	$R \to S$	
4	P	asm, for $\to$I
	$\vdots$	

Solution (Hint 1). Inside the box, what is the current goal, and which available line is a disjunction?

Solution (Hint 2). Line 5 is Q, the assumption of a first case box opened inside the box of P; like the second case box, it must end in the current goal S.

Solution. The moves are two. First, the box of line 4 was read off the goal: $P \to S$ is an implication, so the box opened with P, and its current goal is S; it will be closed by $\to$I. Second, inside it, the available disjunction $Q \vee R$ (line 1, in no box, so available inside the

box) is split, and **both case boxes end in the current goal** S. In the case Q, line 2 fires to give $P \to S$, which fires on line 4; in the case R, line 3 fires.

1	$Q \vee R$	
2	$Q \to (P \to S)$	
3	$R \to S$	
4	$\quad P$	asm, for $\to$I
5	$\quad\quad Q$	asm, for $\vee$E
6	$\quad\quad P \to S$	$\to$E 2, 5
7	$\quad\quad S$	$\to$E 6, 4
8	$\quad\quad R$	asm, for $\vee$E
9	$\quad\quad S$	$\to$E 3, 8
10	$\quad S$	$\vee$E 1, 5–7, 8–9
11	$P \to S$	$\to$I 4–10

Line 7 cites line 4, the assumption of the outer box, which contains line 7. Line 10 is a line of the box 4–10 itself (the case boxes have closed), so $\to$I may cite that box at line 11. Splitting outside the box instead, with both case boxes ending in $P \to S$, is also a correct derivation of the sequent; it does not complete the lines given, because it would split before line 4, not after.

1.6.8.4 R. Review

Exercise. R1

Let f be a function and $[a, b]$ an interval. Take the atomic statements C: "f is continuous on $[a, b]$", B: "f is bounded on $[a, b]$" and I: "f is integrable on $[a, b]$"; you need not know what the words mean. Translate the argument "If f is continuous on $[a, b]$ and bounded there, then f is integrable on $[a, b]$. f is bounded on $[a, b]$. f is continuous on $[a, b]$. So f is integrable on $[a, b]$ and bounded there." into a sequent, and derive it.

Solution. The sequent is $(C \wedge B) \to I,\ B,\ C \vdash I \wedge B$: three premises and a conclusion, as in Definition 1.2.4. The goal is a conjunction, so it splits into the subgoals I and B. B is a premise. The only line that mentions I is the first premise, whose hypothesis $C \wedge B$ is built by $\wedge$I (Definition 1.2.5) from two premises; then $\to$E (Definition 1.3.1) gives I.

1	$(C \wedge B) \to I$	
2	B	
3	C	
4	$C \wedge B$	$\wedge$I 3, 2
5	I	$\to$E 1, 4
6	$I \wedge B$	$\wedge$I 5, 2

Line 4 writes $C \wedge B$, with C on the left as in the hypothesis of line 1 ($\wedge$I may cite 3, 2 or 2, 3); the hypothesis must match exactly. Line 2 is cited twice, at lines 4 and 6: a line is not used up by being cited.

Exercise. R2

Show that $(P \to Q) \wedge (Q \to R),\ P \wedge S \vdash R \wedge S$, and say in one sentence which step you found by working backwards from the goal.

Solution. Premise-directed, both premises are conjunctions and are taken apart by $\wedge$E (Definition 1.2.5); goal-directed, the goal $R \wedge S$ splits into R and S, and S is already in hand. For R, work backwards as in Example (Chains of implications): $Q \to R$ is available, so Q becomes the subgoal; $P \to Q$ is available, so P does; and P is in hand.

1	$(P \to Q) \wedge (Q \to R)$	
2	$P \wedge S$	
3	$P \to Q$	$\wedge$E 1
4	$Q \to R$	$\wedge$E 1
5	P	$\wedge$E 2
6	S	$\wedge$E 2
7	Q	$\to$E 3, 5
8	R	$\to$E 4, 7
9	$R \wedge S$	$\wedge$I 8, 6

The chain of lines 7 and 8 was found backwards, from the subgoal R to Q to P, and written forwards.

Exercise. R3

Is $\vdash (P \wedge (P \to Q)) \to (Q \vee R)$? Justify your answer.

Solution. Yes. The goal's main connective is the outer $\to$, so the move is to open a box with its hypothesis $P \wedge (P \to Q)$ (Definition 1.4.1). Inside, the current goal is $Q \vee R$, which we

postpone. The assumption is a conjunction, so $\wedge$E gives P and $P \to Q$; $\to$E gives Q; now the left side of the goal is in hand, and $\vee$I (Definition 1.5.1) takes it.

1	$P \wedge (P \to Q)$	asm, for $\to$I
2	P	$\wedge$E 1
3	$P \to Q$	$\wedge$E 1
4	Q	$\to$E 3, 2
5	$Q \vee R$	$\vee$I 4
6	$(P \wedge (P \to Q)) \to (Q \vee R)$	$\to$I 1–5

The derivation has no premises, so it shows $\vdash (P \wedge (P \to Q)) \to (Q \vee R)$. The look-alike start fails: the statement contains $\vee$, and one might begin with $\vee$I on the whole goal. But the goal's main connective is $\to$, not $\vee$, so $\vee$I cannot produce it; and even inside the box, choosing the side R of $Q \vee R$ would have left the subgoal R, which nothing supplies; the side Q is the one lines 2–4 supply.

Chapter 2

Negation and Classical Logic

Is there a rational number whose square is 2? Is x^2+1 ever 0 for a real number x? Is "if $x^2 < 0$, then $x = 7$" true? Chapter 1 could prove that things are so, but not that something is **not** so, and it read the third sentence only informally, as a promise not broken: no rule yet derives it. By the end of this chapter you will be able to prove that something is not so; you will know when "suppose not" is a different kind of step from the others, and why that one step is all that classical logic adds. Two questions stay open at the end. How do you find a derivation when none comes to mind? Chapter 3 answers it. And how could you ever show that something is **not** derivable? Chapter 4 answers that.

Steps about "every" and "some". Proofs in this chapter apply facts about every integer to one integer, pass from an arbitrary rational number to all of them, and name the k in "$n = 2k+1$". These are steps about "every" and "some". Here they are used as ordinary reasoning, as in school algebra; Chapter 5 gives them rules.

What you need. All of Chapter 1: the rules for "and", "implies", and "or", and above all the definition of a derivation, with its boxes and its scope rule (Definition 1.4.2). School algebra about integers and real numbers is used in examples, as facts.

Time. About 30 hours on the core track (all of A, one B and one C per section, the exit checks); about 50 hours for everything.

Minimum before *Sets, Relations, and Cardinality*. Chapters 1–6 and 8, core; Sections 7.1–7.2 before its Chapter 3.

2.0.1 Roadmap

- **Contradiction: Using a Negation.** The rules ¬E and ⊥E: what a clash is, what it buys, why an impossible case closes itself and an implication with an impossible hypothesis is true, and the two sequents called the paradoxes of material implication.
- **Proving a Negation.** The rule ¬I, and with it the irrationality of $\sqrt{2}$, then of $\sqrt{p}$ for a prime p.
- **Derived Rules: Proved Sequents as New Rules.** A proved sequent used as one line; modus tollens, the contrapositive and disjunctive syllogism; and why pasting a proved sequent into a derivation is safe (weakening and cut).
- **Iff, Equivalence, and What a Definition Says.** The rules for ↔, provable equivalence, why a definition's "if" means "if and only if", chains of equivalences, and why no statement is equivalent to its own negation.
- **Proof by Contradiction and the Classical Rules.** The rule RAA, the classical laws

it gives (double-negation elimination, excluded middle, De Morgan, Peirce), where the classical step is needed, and why three of its forms stand or fall together.

2.0.2 Named moves

- **Goal $\bot$: find a clash.** Look for φ and $\neg\varphi$ among the available lines, or make φ a subgoal for some $\neg\varphi$ you have. (Section 2.1)
- **An impossible case closes itself.** In a proof by cases, a case that reaches $\bot$ ends by $\bot$E on the current goal. (Section 2.1)
- **To prove "not φ", assume φ.** The box opens with the unnegated statement and ends on $\bot$. (Section 2.2)
- **Use a derived rule as one line.** A proved sequent is a new rule. (Section 2.3)
- **Two boxes for an iff.** One box for each direction, each ending on the other side. (Section 2.4)
- **When stuck, assume the negation of the goal.** Proof by contradiction, and then hunt for a clash. (Section 2.5)
- **A disjunction with neither side within reach: expect the classical step.** When each side of a disjunctive goal breaks on its own, aim at the disjunction by RAA. (Section 2.5)

Section 3.1 gathers these moves and Chapter 1's into one table, and all thirteen rules onto one card.

2.0.3 Summary

The chapter added the rules for $\neg$, $\bot$ and $\leftrightarrow$ and the rule RAA, which with Chapter 1's seven make the thirteen primitive rules. It proved that a clash gives anything (Proposition 2.1.1) and that an impossible hypothesis makes an implication true (Proposition 2.1.2); it proved the theorem Argument B set out to prove (Theorem 2.1.3) and that $\sqrt{2}$ is irrational (Theorem 2.2.1). It turned proved sequents into rules, with weakening and cut as the reason pasting is safe (Lemma 2.3.3). It showed that an iff is two implications, that a definition's "if" is an iff, and that no statement is equivalent to its own negation (Lemma 2.4.2). With RAA it derived double negation, excluded middle, De Morgan's laws and Peirce's law, and showed RAA, $\neg\neg$E and LEM to be one principle (Theorem 2.5.7). Still open: a method for **finding** derivations (Chapter 3), and a way to show that a sequent has **no** derivation (Chapter 4), which also settles whether the classical rule could be done without.

2.0.4 Terms

- **biconditional rules** $\leftrightarrow$I, $\leftrightarrow$E: prove by two boxes; use either direction (Definition 2.4.1).
- **classical rules** RAA, $\neg\neg$E, LEM: interderivable; one principle (Theorem 2.5.7).
- **contradiction** $\bot$: the line a clash produces (Definition 2.1.1).
- **contradiction rules** $\neg$E, $\bot$E: a clash gives $\bot$; $\bot$ gives anything (Definition 2.1.1).
- **contrapositive**: $\neg\psi \to \neg\varphi$, derivable from $\varphi \to \psi$ (Proposition 2.3.2).
- **cut**: a proved lemma may be pasted in (Lemma 2.3.3).
- **De Morgan's laws**: how negation passes through "and" and "or" (Proposition 2.5.5).
- **derived rule**: a proved sequent used as one line (Definition 2.3.1).
- **disjunctive syllogism** DS: from $\varphi \vee \psi$ and $\neg\varphi$, conclude ψ (Proposition 2.3.4).
- **double negation**: $\neg\neg\varphi$ and φ, each from the other (Proposition 2.5.1).
- **ex falso**: from a clash, anything follows (Proposition 2.1.1).
- **excluded middle** LEM: $\varphi \vee \neg\varphi$, with no premises (Proposition 2.5.3).
- **modus tollens** MT: from $\varphi \to \psi$ and $\neg\psi$, conclude $\neg\varphi$ (Proposition 2.3.1).

- **negation rule** ¬I: assume φ, reach $\bot$, conclude $\neg\varphi$ (Definition 2.2.1).
- **Peirce's law**: $((\varphi \to \psi) \to \varphi) \to \varphi$ (Proposition 2.5.4).
- **provably equivalent**, $\dashv\vdash$: each derivable from the other (Definition 2.4.2).
- **reductio ad absurdum** RAA: assume $\neg\varphi$, reach $\bot$, conclude φ (Definition 2.5.1).
- **vacuously true**: an implication whose hypothesis cannot hold (Proposition 2.1.2).
- **weakening**: extra premises never spoil a derivation (Lemma 2.3.3).

2.1 Contradiction: Using a Negation

Chapter 1 gave rules for "and", "implies", and "or", and not one of them mentions "not". Yet written proofs are full of sentences such as "this is impossible", "this case cannot occur", and "which contradicts what we assumed". This section gives the two rules that **use** a negation: the first says exactly what a clash between two lines is, and the second says what a clash buys. The rule that **proves** a negation is the subject of the next section.

2.1.1 A case that cannot occur

Here is Argument A of Section 1.1, told as a proof by cases. Let n be an integer whose square is even. Every integer is even or odd, so there are two cases. If n is even, we are done. If n is odd, then n^2 is odd; but n^2 is even, and no integer is both, so this case cannot occur. Hence n is even.

Chapter 1 has a rule for a proof by cases: $\vee$E of Definition 1.6.1. It asks for a disjunction and two boxes, one opened by each disjunct, and **both boxes must end on the same line**, the goal. In the argument above the goal is "n is even". The first box reaches it at once. The second box does not reach it at all. It reaches "impossible", and then simply declares itself finished.

Two things are missing, and no rule of Chapter 1 supplies them.

- A way to **record** the impossibility: the lines "n^2 is even" and "n^2 is odd", together with "no integer is both", should produce a line that says "contradiction".
- A way to **close** the impossible box with the goal: from the line "contradiction", the box should be allowed to write "n is even", so that both boxes end on the same line and $\vee$E applies.

Section 1.1 introduced the symbols for both. The connective $\neg$ forms the **negation** $\neg\varphi$, read "not φ". The constant $\bot$, read "contradiction", is a statement that is false however things stand. There they were symbols only. The two rules below give them their first uses.

2.1.2 The rules for a clash

The rule for recording a clash has a simple content. The statement $\neg\varphi$ says that φ fails, so a derivation that has both φ and $\neg\varphi$ available is working in a situation that cannot occur. The rule writes this down as the line $\bot$. A second rule then says what such a line is good for.

A statement and its negation clash; from a clash, anything follows.

Definition 2.1.1. Rules for Contradiction

i	φ			i	$\bot$	
j	$\neg\varphi$			k	φ	$\bot$E i
k	$\bot$	$\neg$E i, j				

The left schema is ¬E (**not-elimination**): if φ and $\neg\varphi$ stand on lines i and j available at line k, write $\bot$, citing both, the unnegated line first. The right schema is $\bot$E (**contradiction-elimination**, or *ex falso*): if $\bot$ stands on a line i available at line k, write **any** statement φ, citing line i.

Why may $\bot$E write anything? A line $\bot$ can stand only where the premises and open assumptions cannot all hold together: a derivation that reaches $\bot$ describes a situation that never happens. Nothing said there can ever be caught false, just as a promise whose condition is never met is never broken; the subsection "Vacuous truth" below returns to this.

As in every schema (Definition 1.2.1), φ is one statement, the same wherever it occurs, and i, j, k are line numbers. We read ¬E in four bullets.

- **Two given lines,** both available (the scope rule of Definition 1.4.2), in either order on the page.
- **Exactly φ and $\neg\varphi$:** the second is the first with one ¬ in front of the whole of it, symbol for symbol. "Contradicts it in spirit" is not enough: $x < 1$ and $x > 1$ cannot both hold, but neither is the other's negation as written (the box below).
- **The citation** lists the unnegated line first: "¬E i, j" says that line i is φ.
- **The new line** is $\bot$ and nothing else; what to do with the clash is the business of $\bot$E.

And $\bot$E in three.

- **Schema.** The right one in the block above.
- **Reading.** One available line $\bot$, cited alone; write **any** statement, related to nothing above, a negation or $\bot$ itself included.
- **Example.** From line 4 $\bot$, "$\bot$E 4" writes R, or $\neg R$, or $x = 7$.

Second reading · *Where a line $\bot$ comes from, and why $\bot$E is honest*

Two more observations complete the reading. First, $\bot$E is the **elimination** rule of $\bot$: it says how a line $\bot$ is used. There is no introduction rule of its own for $\bot$. Other rules can pass a $\bot$ along ($\bot$E may write it again, and $\to$E writes it from $\varphi \to \bot$ and φ), but each of them (R, $\wedge$E, $\to$E, $\vee$E, whenever their new line is $\bot$) needs $\bot$ already written, perhaps inside a longer statement, in a line it uses ($\vee$I can bring $\bot$ in, but only inside a disjunction). Unless $\bot$ already occurs in a premise or an assumption, the first line of a derivation that is $\bot$ itself is written by ¬E: it is the one rule whose new line is always $\bot$, from lines in which $\bot$ need not occur. So ¬E plays two parts at once: it is the way a negation is used, and it is the way $\bot$ is produced.

Second, why is $\bot$E honest? A rule should never lead from lines that hold to a line that fails. A line $\bot$ can stand only in a part of a derivation where the premises and the open assumptions cannot all hold together, at least when the $\bot$ was written by ¬E, from some φ and $\neg\varphi$ that were both available there; that the rules which only pass a $\bot$ along keep this property is part of what Section 1.2's Looking-ahead box promised for every rule. In such a place there is no situation in which the conclusion of $\bot$E would have to be checked, so it cannot be caught failing.

In the Givens/Goal table, ¬E adds $\bot$ to the Givens once φ and $\neg\varphi$ are both there; and with $\bot$ among the Givens, $\bot$E ticks off any goal.

Quick check. Which of these pairs of lines are, as written, an instance of the given lines of ¬E, and in which order must each be cited? The pairs are (i) $Q \wedge R$ and $\neg(Q \wedge R)$; (ii) $\neg P$ and $\neg\neg P$; (iii) $P \to Q$ and $\neg P \to Q$.

Solution.

(i) Yes, with $\varphi = Q \wedge R$; cite the line $Q \wedge R$ first. (ii) Yes, with $\varphi = \neg P$: the second line is $\neg(\neg P)$. Cite $\neg P$ first, then $\neg\neg P$. (iii) No. The main connective of $\neg P \to Q$ is $\to$ (Definition 1.1.2), so it is not a negation at all; the $\neg$ applies to P alone. The negation of $P \to Q$ would be $\neg(P \to Q)$.

Remark. **Convention: negations of order and equality.** For real numbers a and b, exactly one of $a < b$ and $a \geq b$ holds. So $a \geq b$ says the same as $\neg(a < b)$, and likewise $a > b$ the same as $\neg(a \leq b)$, $a \leq b$ as $\neg(a > b)$, and $a < b$ as $\neg(a \geq b)$; and $a \neq b$ **means** $\neg(a = b)$, by definition. A derivation writes the negation out whenever a rule for "not" must see it (¬E here, and the rule of the next section); otherwise, and in prose, the short forms stay. Together with the box of Section 1.5, which writes $a \leq b$ out as $(a < b) \vee (a = b)$ when $\vee$I or $\vee$E must see the disjunction, this fixes how an inequality looks on a line.

The fact "no integer is both even and odd" is a negation of a conjunction, $\neg(E \wedge O)$ with E for "n is even" and O for "n is odd". It clashes only with a line $E \wedge O$, which $\wedge$I builds from separate lines E and O.

Quick check. For a real number x, which pairs are a clash as written, once the convention is used? (i) $x^2 < 0$ and $x^2 \geq 0$; (ii) $x = 2$ and $x \neq 2$; (iii) $x < 1$ and $x > 1$.

Solution.

(i) Yes: $x^2 \geq 0$ is $\neg(x^2 < 0)$; cite $x^2 < 0$ first. (ii) Yes: $x \neq 2$ is $\neg(x = 2)$; cite $x = 2$ first. (iii) No: $x > 1$ is $\neg(x \leq 1)$, and $x < 1$ is not $x \leq 1$. The line $x \leq 1$, written out, must first be built by $\vee$I (drill A2(l)).

Warning. A clash must have the exact shape φ, $\neg\varphi$. And some pairs that look opposed do not clash at all: $x \geq 1$ and $x \leq 1$ both hold at $x = 1$.

Warning. $\bot$E does not run backwards. From $\bot$ we may write φ, or $\neg\varphi$, or anything; but from a line φ alone no rule writes $\bot$, unless $\bot$ already occurs in it. If it could, the true line $x = 1$ would give $\bot$, and then $\bot$E would give the false line $x = 2$.

2.1.3 First derivations with a clash

The five examples below grow in difficulty. Each Idea says why each rule was chosen, in the two directions Chapter 1 used: from the goal (what does the goal's main connective ask for?) and from the premises (what does an available line's main connective let us do?). This section adds one question to each list. From the goal: *when nothing produces the goal, can a clash produce it?* From the premises: *an available negation $\neg\varphi$ is used in exactly one way, by $\neg$E, so it asks for φ.*

Example. The first clashes

Show that (a) $\bot \vdash P$, and (b) $P \wedge Q,\ \neg Q \vdash R$.

Idea.

(a) is the rule $\bot$E itself, the degenerate case with no clash to find. In (b), the first premise is a conjunction, so the premise-directed move takes it apart by $\wedge$E. Its half Q and the premise $\neg Q$ are exactly a pair φ, $\neg\varphi$. The goal R has nothing to do with the premises: no rule builds it from them. Only $\bot$E can write it, and $\bot$E needs a line $\bot$, which the clash supplies.

Snapshots for (b): the goal alone; then the goal handed to $\bot$E, with $\bot$ as the new goal.

1	$P \wedge Q$
2	$\neg Q$
	⋮
5	R

1	$P \wedge Q$	
2	$\neg Q$	
	⋮	
4	$\bot$	
5	R	$\bot$E 4

Solution.

(a) Two lines: the premise, and one application of $\bot$E.

1	$\bot$	
2	P	$\bot$E 1

(b) Five lines.

1	$P \wedge Q$	
2	$\neg Q$	
3	Q	$\wedge$E 1
4	$\bot$	$\neg$E 3, 2
5	R	$\bot$E 4

Line 4 cites line 3 first because line 3 plays φ and line 2 plays $\neg\varphi$, with $\varphi = Q$. The half P of the first premise is never used; a derivation need not use every premise (Definition 1.2.3).

The next result is the pattern of (b) with nothing extra around it. We state it as a proposition, because later sections cite it.

Proposition 2.1.1. Contradictory Premises Prove Anything

$P,\ \neg P \vdash Q$.

Idea. The premises are a clash as they stand, so ¬E writes $\bot$ at once, and $\bot$E writes the goal, which nothing else could produce.

Proof. The derivation

1	P	
2	$\neg P$	
3	$\bot$	¬E 1, 2
4	Q	$\bot$E 3

has premises P and $\neg P$ and last line Q. This proves the proposition.

■

In the next derivation the clash happens inside a box, and the rule $\bot$E writes the box's goal.

Proposition 2.1.2. Vacuous Implication

$\neg P \vdash P \to Q$.

Idea. The goal is an implication, so the goal-directed move of Chapter 1 ($\to$I, Definition 1.4.1) opens a box with its hypothesis P; the new goal is Q. Nothing available mentions Q, so no rule can build it. But inside the box, the assumption P and the premise $\neg P$ clash, and $\bot$E writes the box's goal Q. This is the section's goal-directed move: *when nothing produces the goal, look for a clash.*

Snapshots: the box with its goal Q; then the clash named as the subgoal.

1	$\neg P$	
2	$\quad P$	asm, for $\to$I
	$\quad \vdots$	
4	$\quad Q$	
5	$P \to Q$	$\to$I 2–4

1	$\neg P$	
2	$\quad P$	asm, for $\to$I
	$\quad \vdots$	
3	$\quad \bot$	
4	$\quad Q$	$\bot$E 3
5	$P \to Q$	$\to$I 2–4

Proof. The derivation

1	$\neg P$	
2	$\quad P$	asm, for $\to$I
3	$\quad \bot$	$\neg$E 2, 1
4	$\quad Q$	$\bot$E 3
5	$P \to Q$	$\to$I 2–4

has premise $\neg P$ and last line $P \to Q$. Line 3 may cite line 1, which lies in no box. This proves the proposition.

■

Line 4 is the line that makes the box close properly: $\to$I needs the box to end on Q, as a line of that box itself, and $\bot$E put it there.

In the next example no clash is lying about. It has to be made.

Example. Manufacturing the other half of a clash

Show that $\neg(P \vee Q) \vdash P \to R$.

Idea. As before, the goal $P \to R$ opens a box on P, and the new goal R occurs nowhere else, so a clash must write it. Inside the box there is no $\neg P$ to clash with. The only negation available is the premise $\neg(P \vee Q)$, and the premise-directed question says what it asks for: the line $P \vee Q$. That line is one $\vee$I away from the assumption P. So the clash is **manufactured**: the unnegated half of the premise becomes a subgoal, and $\vee$I is chosen not because the goal is a disjunction but because the premise's clash needs one.

Snapshots: the box with its goal R; then the subgoal $\bot$, for which the premise asks for $P \vee Q$.

1	$\neg(P \vee Q)$	
2	$\quad P$	asm, for $\to$I
	$\quad \vdots$	
5	$\quad R$	
6	$P \to R$	$\to$I 2–5

1	$\neg(P \vee Q)$	
2	$\quad P$	asm, for $\to$I
	$\quad \vdots$	
4	$\quad \bot$	
5	$\quad R$	$\bot$E 4
6	$P \to R$	$\to$I 2–5

Solution. Six lines.

1	$\neg(P \vee Q)$	
2	$\quad P$	asm, for $\to$I
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad \bot$	$\neg$E 3, 1
5	$\quad R$	$\bot$E 4
6	$P \to R$	$\to$I 2–5

Line 4 cites line 3 as φ and line 1 as $\neg\varphi$, with $\varphi = P \vee Q$. In words: if neither P nor Q holds, then P cannot hold, so an implication with hypothesis P is true whatever its conclusion. The mirror image, with Q as the hypothesis, is exercise B1.

Quick check. Inside the box of P, could $\bot$E write R straight from line 2? Which lines must come first?

Solution. No. $\bot$E needs a line $\bot$, and $\bot$ needs the pair $P \vee Q$, $\neg(P \vee Q)$: first $P \vee Q$ by $\vee$I 2, then $\bot$ by $\neg$E 3, 1.

The last of these first derivations is the pattern of the opening argument: a case split in which one case is impossible.

Example. An impossible case

Show that $P \vee Q,\ P \to R,\ \neg R \vdash Q$.

Idea. The goal Q is a letter, so the goal gives no introduction rule to aim at. The premises do: $P \vee Q$ is a disjunction, and Chapter 1's premise-directed move for a disjunction is $\vee$E, with two boxes that both end in the current goal Q. The box of Q ends as soon as it opens: its assumption is already the goal. The box of P cannot reach Q honestly, because nothing in it produces Q. So we look for a clash: P feeds the premise $P \to R$, and R clashes with $\neg R$. The box ends on Q by $\bot$E. This is the second new move: *an impossible case closes*

itself.

Snapshots: the split set up; then the first box's goal handed to ⊥E.

1	$P \vee Q$	
2	$P \to R$	
3	$\neg R$	
4	$\quad P$	asm, for ∨E
	$\quad \vdots$	
7	$\quad Q$	
8	$\quad Q$	asm, for ∨E
9	Q	∨E 1, 4–7, 8–8

1	$P \vee Q$	
2	$P \to R$	
3	$\neg R$	
4	$\quad P$	asm, for ∨E
	$\quad \vdots$	
6	$\quad \bot$	
7	$\quad Q$	⊥E 6
8	$\quad Q$	asm, for ∨E
9	Q	∨E 1, 4–7, 8–8

Solution. Nine lines.

1	$P \vee Q$	
2	$P \to R$	
3	$\neg R$	
4	$\quad P$	asm, for ∨E
5	$\quad R$	→E 2, 4
6	$\quad \bot$	¬E 5, 3
7	$\quad Q$	⊥E 6
8	$\quad Q$	asm, for ∨E
9	Q	∨E 1, 4–7, 8–8

The second box consists of its assumption alone, a one-line box, just as in the derivation of $P \vee P \vdash P$ in Example (First derivations with "or"): its last line is its first line, and that line is Q, which is what ∨E needs. The two boxes are separate: nothing in the second box could cite lines 4–7.

Look at line 7 of this derivation again. The box of P is a case "that cannot occur", and line 7 is the sentence "so there is nothing to prove": the case is closed with whatever the other case reaches. Every proof by cases with an impossible case has this line somewhere, usually unwritten.

Warning. The impossible case must be impossible **from what is available in its own box**. A clash with a line of a sibling box, or of a box that has already closed, is a scope error (Definition 1.4.2), even when the clash "looks right". In Example (An impossible case) the clash in the box of P uses line 5, from that box, and the premise on line 3; neither lives in

another box. Exercise C2 shows the error and what it can prove.

The moves of this section, in one place. **Goal $\bot$: find a clash.** Look for φ and $\neg\varphi$ among the available lines; if only $\neg\varphi$ is there, make φ a subgoal. **An impossible case closes itself.** In a case split, a case that reaches $\bot$ ends by $\bot$E on the current goal. The next section adds a way to reach the goal $\bot$ deliberately, by assuming something.

2.1.4 How the rules look in a mathematical proof

We can now prove the fact of Argument A in Section 1.1, with every rule named. Three school facts about integers enter as premises. An integer n is **even** if $n = 2k$ for some integer k, and **odd** if $n = 2k + 1$ for some integer k. And:

(E1) every integer is even or odd;

(E2) no integer is both even and odd;

(E3) the square of an odd integer is odd.

The third is a computation. If $n = 2k + 1$ for an integer k, then

$$n^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1,$$

and $2k^2 + 2k$ is an integer, so n^2 is odd.

Theorem 2.1.3. Even Square, Even Number

Let n be an integer. If n^2 is even, then n is even.

Idea. The hypothesis "n^2 is even" hands us nothing to compute with: it says $n^2 = 2j$ for some integer j, and no algebra turns that into a statement about n. The word "odd" does hand us something, since $n = 2k + 1$ can be squared. So we split on (E1) and show that the odd case is impossible: there (E3) makes n^2 odd, which clashes with the hypothesis through (E2). The impossible case then closes itself.

Proof. Suppose n^2 is even [assumption for $\to$I]. By (E1), n is even or n is odd, and we argue by cases [$\vee$E].

Case n even. This is the goal.

Case n odd. By (E3), n^2 is odd [$\to$E]. So n^2 is both even and odd [$\wedge$I]; this contradicts (E2) [$\neg$E]. So this case cannot occur, and in it n is even [$\bot$E].

In both cases n is even [$\vee$E]. This proves that if n^2 is even, then n is even [$\to$I].

■

The logic of the proof fits in a derivation. Write E for "n is even", O for "n is odd", S for "n^2 is even", and T for "n^2 is odd". Then (E1) for n is $E \vee O$, (E2) for n^2 is $\neg(S \wedge T)$, and (E3) for n is $O \to T$; these are the premises.

1	$E \vee O$	
2	$\neg(S \wedge T)$	
3	$O \to T$	
4	$\quad S$	asm, for $\to$I
5	$\quad\quad E$	asm, for $\vee$E
6	$\quad\quad O$	asm, for $\vee$E
7	$\quad\quad T$	$\to$E 3, 6
8	$\quad\quad S \wedge T$	$\wedge$I 4, 7
9	$\quad\quad \bot$	$\neg$E 8, 2
10	$\quad\quad E$	$\bot$E 9
11	$\quad E$	$\vee$E 1, 5–5, 6–10
12	$S \to E$	$\to$I 4–11

Each line is one sentence of the proof.

Line	Sentence of the proof
1–3	(E1) for n, (E2) for n^2, (E3) for n
4	Suppose n^2 is even.
5	Case n even. (A one-line box: it is the goal.)
6	Case n odd.
7	By (E3), n^2 is odd.
8	So n^2 is both even and odd,
9	which contradicts (E2).
10	So this case cannot occur, and in it n is even.
11	In both cases n is even.
12	This proves that if n^2 is even, then n is even.

Line 8 cites line 4, the assumption of the outer box; it may, since the box 4–11 contains line 8 (Definition 1.4.2). Two steps of the proof are not in the derivation: passing from the facts (E1)–(E3), about every integer, to their instances for this n and this n^2, and naming the k of (E3). Both are steps about "every" and "some", used as ordinary reasoning (the note at the start of this chapter).

Quick check. Which sentence of the proof of Theorem 2.1.3 is the rule $\bot$E, and what does it write?

Solution. The sentence "So this case cannot occur, and in it n is even", line 10 of the derivation. It cites the line $\bot$ (line 9) and writes E, "n is even", which is the goal of the case, so that the box of O ends on the same line as the box of E.

Every rule in this proof is a primitive rule: $\to$I, $\vee$E, $\to$E, $\wedge$I, $\neg$E, and $\bot$E. The next section uses the theorem twice, to prove that no rational number has square 2. Section 2.3 reads the same proof again with shorter steps, as mathematicians usually write it, "by the contrapositive".

2.1.5 Vacuous truth

We can now settle the question that makes many readers distrust logic. Is the sentence

> if $x^2 < 0$, then $x = 7$

true, for a real number x? The hypothesis never holds, and the conclusion has nothing to do with it. Everyday intuition is unsure what to say. The rules are not unsure.

Example. An implication with an impossible hypothesis

Let x be a real number. Show that if $x^2 < 0$, then $x = 7$.

Idea. The goal is an implication, so open a box with $x^2 < 0$, aiming at $x = 7$. Nothing produces $x = 7$, so look for a clash: the school fact $x^2 \geq 0$ is $\neg(x^2 < 0)$, which clashes with the assumption. This is exactly the derivation of Proposition 2.1.2, with P read as "$x^2 < 0$" and Q as "$x = 7$".

Solution. With the school fact as the premise:

1	$\neg(x^2 < 0)$	
2	$\quad x^2 < 0$	asm, for $\to$I
3	$\quad \bot$	$\neg$E 2, 1
4	$\quad x = 7$	$\bot$E 3
5	$x^2 < 0 \to x = 7$	$\to$I 2–4

In prose: *suppose $x^2 < 0$; this contradicts $x^2 \geq 0$; so this case cannot occur, and in it $x = 7$; hence if $x^2 < 0$, then $x = 7$.*

We wrote out all five lines rather than citing Proposition 2.1.2 as a single step. Using an already proved sequent as one line of a new derivation is a genuine new permission, and Section 2.3 grants it; until then, a derivation repeats the steps.

An implication whose hypothesis cannot hold is called **vacuously true**. Nothing about it was decided by convention. It is forced by two rules that each look harmless: $\to$I, which proves an implication by assuming its hypothesis, and $\bot$E, which closes an impossible case.

Everyday language agrees. A friend says: "if you finish the exercises by Friday, I will buy you lunch". You do not finish them. Whatever your friend does on Friday, the promise has not been broken. *A promise never tested is never broken.*

Vacuous truth is not a curiosity; it is what makes case splits work. In Example (An impossible case), line 7 closed the impossible case by $\bot$E, and line 10 of the derivation of Theorem 2.1.3

did the same.

Warning. "True" is not "useful". The statements "if $x^2 < 0$, then $x = 7$" and "if $x^2 < 0$, then $x \neq 7$" are both true, and neither says anything about x or about 7. Likewise "if n is an even prime greater than 2, then $n = 17$" is true and empty. When a hypothesis turns out to be impossible, suspect the hypothesis; do not believe the conclusion.

2.1.6 Two sequents that sound wrong

Two sequents in hand sound wrong when read aloud with unrelated statements. The first is $Q \vdash P \to Q$: a box opened with P, whose only other line repeats Q (line 1 lies in no box), closed by $\to$I, as in Exercise B2 of Section 1.4. The second is $\neg P \vdash P \to Q$, Proposition 2.1.2. Read aloud: "7 is odd; so, if every square is negative, then 7 is odd", and "7 is not even; so, if 7 is even, then every square is negative". They are called the **paradoxes of material implication**.

Each comes from a rule, so whoever rejects one must say which rule goes: the first rests on $\to$I closing a box whose assumption no line uses (with R copying a line in from outside), the second on $\bot$E. Some logics, built for purposes other than mathematics, do drop $\bot$E; this book keeps it, and Exercise B3 of Section 2.5 shows that the classical rule would bring it back anyway. The same rule gives Proposition 2.1.1: from a clash, anything.

Mathematics does not suffer: its "if" almost always sits under "every", joining two statements about one variable, as in "for every integer n, if n is a multiple of 4, then n is even". Instances with a false hypothesis ("if 3 is a multiple of 4, then 3 is even") are true and empty, as the paradoxes say; the content lies where the hypothesis holds, and there the promise is exact. Nobody wants to prove a sentence whose two parts share nothing.

Quick check. Let x be a real number. Which of the two sequents makes "if $x^2 < 0$, then $x > 5$" true, and which makes "if $x < -3$, then $x^2 + 1 > 0$" true?

Solution. The first by $\neg P \vdash P \to Q$, since $x^2 < 0$ fails ($x^2 \geq 0$ is $\neg(x^2 < 0)$, the convention box); the second by $Q \vdash P \to Q$, since $x^2 + 1 > 0$ holds whatever x is.

Second reading · *The 'if' of a counterfactual*

"If 7 were even, 7 would be divisible by 2" sounds true, and "if 7 were even, 7 would be divisible by 3" sounds false, though the two share the false hypothesis "7 is even". By $\neg P \vdash P \to Q$, both material implications are true. So the subjunctive "if" of everyday talk is not the connective $\to$: whether it holds depends on more than the truth of its two parts. Mathematics rarely uses it. When a proof says "if n were odd, then n^2 would be odd", it is working inside a box whose assumption is "n is odd", and its "if" is the material one.

2.1.7 The rule card

Before the card, the two rules at work on large statements, one use per row.

Example. Instances of ¬E and ⊥E

Rule	φ	Cited lines	New line
¬E	$P \vee (Q \wedge R)$	$P \vee (Q \wedge R)$, then $\neg(P \vee (Q \wedge R))$	$\bot$
¬E	$\neg(P \to Q)$	$\neg(P \to Q)$, then $\neg\neg(P \to Q)$	$\bot$
⊥E	$(P \wedge Q) \to \neg R$	$\bot$	$(P \wedge Q) \to \neg R$

Not instances. (i) The lines $P \to Q$ and $P \to \neg Q$ are no pair for ¬E: the second is an implication, not $\neg(P \to Q)$, and no single φ makes the two lines φ and $\neg\varphi$. (ii) From the line $P \wedge \neg P$, "Q — ⊥E" is not an instance: ⊥E cites a line that is $\bot$ itself. Take P and $\neg P$ out by ∧E, write $\bot$ by ¬E, then ⊥E.

Chapter 1's seven rules are on the card at the end of Section 1.6; here are this chapter's so far.

Rule	Shape	You write	Lean
¬E	$\varphi, \neg\varphi \Longrightarrow \bot$	"..., which contradicts ..."	`exact hn hp` (goal `False`)
⊥E	$\bot \Longrightarrow$ any φ	"This case cannot occur." "So there is nothing to prove."	`exfalso`

In Lean the two steps are often taken together: `absurd hp hn` closes any goal from a proof `hp` of P and a proof `hn` of $\neg P$, which is ¬E followed by ⊥E, and the tactic `contradiction` finds such a pair by itself.

Warning (Not allowed: ¬E and ⊥E). Each step breaks the bullet of the rule's reading named after it.

- "$\bot$ from $x < 1$ and $x > 1$, as written": the lines must be exactly φ and $\neg\varphi$ (*Exactly φ and $\neg\varphi$*; the convention box, drill A2(l)).
- "$\bot$ — ¬E j, i", with line j the negation: the unnegated line is cited first (*The citation*).
- "$\bot$ from the one line P": ¬E needs two lines, and ⊥E runs from $\bot$, never to it (*Two given lines*; the warning that ⊥E does not run backwards).

Remark (Looking ahead). Chapter 4 gives each connective a truth table. The rows of $\to$ with a false hypothesis are exactly the ones that →I and ⊥E force here.

2.1.8 Exercises

2.1.8.1 A. Check your understanding

Exercise. A1

(a) Write the schemas of ¬E and ⊥E from memory.

(b) Match each sentence, taken from proofs, with the rule it uses, ¬E or ⊥E: "But n was assumed odd, and we have just shown it is even"; "In this case $x^2 < 0$, which is impossible, so there is nothing to prove".

(c) True or false, with a reason: ⊥E lets us write ⊥ from any line.

(d) In Example (An impossible case), why is the second box only one line long?

(e) In the derivation after Theorem 2.1.3, why does line 8 need line 4, and why may it cite it?

Solution.

(a) ¬E: from a line i with φ and a line j with $\neg\varphi$, both available, write ⊥ at line k, citing i, j. ⊥E: from an available line i with ⊥, write any statement φ at line k, citing i. The two schemas are those of Definition 2.1.1.

(b) The first is ¬E: "odd" and "even" clash through (E2), and the sentence records the clash. The second is ⊥E: "impossible" records a clash (an ¬E against $x^2 \geq 0$, left unsaid), and "there is nothing to prove" closes the case with whatever the case needs.

(c) False. ⊥E runs from ⊥ to any statement, not from any statement to ⊥. ⊥E needs a line ⊥ to start from. The rule that records a clash as ⊥ is ¬E, and it needs two lines φ and $\neg\varphi$.

(d) The box is opened by Q, and Q is already the goal of both boxes. So the box's first line is also its last line, which is all ∨E asks of it.

(e) Line 8 is $S \wedge T$, built by ∧I from S and T; line 4 is the only line that says S. It may cite line 4 because the only box containing line 4, the box 4–11, also contains line 8 (the scope rule of Definition 1.4.2).

Exercise. A2: Drill

Each part takes a line or two. Lines are numbered as in a derivation. The parts run from one step, through a missing line, a justification column and steps to judge, to short derivations.

(a) Lines 1 $\neg(P \wedge Q)$, 2 $P \wedge Q$. What does "¬E 2, 1" write?

(b) Lines 1 $\neg\neg P$, 2 $\neg P$. Which line is φ for ¬E, and how is the step cited?

(c) Line 1 is ⊥. Which of P, $\neg P$, ⊥, $x = 7$ may "⊥E 1" write?

(d) Lines 4 $(P \to Q) \wedge R$ and 6 $\neg((P \to Q) \wedge R)$ are available at line 7. Write line 7 by ¬E, with its citation.

(e) Line 7 is ⊥, justified "¬E 5, 2", and line 2 is $\neg(Q \vee R)$. What is line 5?

(f) Line 9 is $\neg P \vee R$, justified "⊥E 8". What is line 8?

(g) Justify each line, the first three being premises: 1 $\neg(P \vee Q)$, 2 $R \to P$, 3 R, 4 P, 5 $P \vee Q$, 6 ⊥, 7 S.

(h) Lines 1 $\neg P \wedge Q$, 2 P. Is "3. $\bot$ — $\neg$E 2, 1" legal?

(i) Line 1 is P. May "2. $\bot$ — $\bot$E 1" be written?

(j) For a real number x, do the lines $x \geq 1$ and $x \leq 1$ clash?

(k) Find the illegal line: 1 $P \to Q$; 2 $\neg Q$; 3 | P; 4 | Q, $\to$E 1, 3; 5 | $\bot$, $\neg$E 2, 4; 6 | R, $\bot$E 5; 7 $P \to R$, $\to$I 3–6.

(l) For one real number x, lines 1 $x < 1$ and 2 $x > 1$ are given. Make them a clash, line by line.

(m) Derive $P,\ \neg P \wedge Q \vdash R$.

(n) Derive $\neg Q \vdash Q \to P$.

(o) Derive $P \to Q,\ \neg Q,\ P \vdash R$.

Solution.

(a) $\bot$.

(b) $\varphi = \neg P$ (line 2), and $\neg\varphi = \neg\neg P$ (line 1): the step is cited "$\neg$E 2, 1".

(c) All four: $\bot$E writes any statement.

(d) 7 $\bot$, "$\neg$E 4, 6": line 4 is φ, line 6 is $\neg\varphi$.

(e) $Q \vee R$, the unnegated form of line 2.

(f) $\bot$: $\bot$E cites one line, and it is $\bot$.

(g) 4 $\to$E 2, 3; 5 $\vee$I 4; 6 $\neg$E 5, 1; 7 $\bot$E 6.

(h) No: line 1's main connective is $\wedge$, so it is not a negation. First write $\neg P$ by $\wedge$E 1 as line 3, then $\bot$ by $\neg$E 2, 3.

(i) No: $\bot$E starts from a line $\bot$; no rule writes $\bot$ from a single line P.

(j) No: both hold at $x = 1$.

(k) Line 5: the unnegated line comes first. With line 2 as φ, line 4 would have to be $\neg\neg Q$. Correct: "$\neg$E 4, 2".

(l) Write line 2 as $\neg((x < 1) \vee (x = 1))$ ($x > 1$ says that $x \leq 1$ fails, and $x \leq 1$ is written out because $\vee$I must build it); 3 $(x < 1) \vee (x = 1)$, $\vee$I 1; 4 $\bot$, $\neg$E 3, 2.

(m) 3 $\neg P$, $\wedge$E 2; 4 $\bot$, $\neg$E 1, 3; 5 R, $\bot$E 4.

(n) 2 | Q, assumption; 3 | $\bot$, $\neg$E 2, 1; 4 | P, $\bot$E 3; 5 $Q \to P$, $\to$I 2–4.

(o) 4 Q, $\to$E 1, 3; 5 $\bot$, $\neg$E 4, 2; 6 R, $\bot$E 5.

2.1.8.2 B. Practice

Exercise. B1: Two clashes

(a) Show that $\neg(P \vee Q) \vdash Q \to R$.

(b) Show that $(P \to Q) \wedge \neg Q,\ P \vdash \neg R$, and say why $\bot$E may write a negation.

Solution.

(a) As in Example (Manufacturing the other half of a clash), with Q in place of P: the box of Q manufactures $P \vee Q$ by $\vee$I on the other side.

1	$\neg(P \vee Q)$	
2	$\quad Q$	asm, for $\to$I
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad \bot$	$\neg$E 3, 1
5	$\quad R$	$\bot$E 4
6	$Q \to R$	$\to$I 2–5

(b) Take the first premise apart, fire the implication on the second premise, and clash.

1	$(P \to Q) \wedge \neg Q$	
2	P	
3	$P \to Q$	$\wedge$E 1
4	$\neg Q$	$\wedge$E 1
5	Q	$\to$E 3, 2
6	$\bot$	$\neg$E 5, 4
7	$\neg R$	$\bot$E 6

Line 7 is an instance of $\bot$E with $\varphi = \neg R$: the schema allows **any** statement in place of φ, and a negation is a statement.

Exercise. B2: An even successor

Let n be an integer. Prove that if $n+1$ is even, then n is odd, by cases on (E1) with one case shown impossible. Tag each step with its rule, as in the proof of Theorem 2.1.3.

Solution. Suppose $n+1$ is even [assumption for $\to$I]. By (E1), n is even or n is odd, and we argue by cases [$\vee$E].

Case n even. Write $n = 2k$ with k an integer. Then $n+1 = 2k+1$, so $n+1$ is odd, by the definition of odd [$\to$E]. So $n+1$ is both even and odd [$\wedge$I]; this contradicts (E2) [$\neg$E]. So this case cannot occur, and in it n is odd [$\bot$E].

Case n odd. This is the goal.

In both cases n is odd [$\vee$E]. This proves that if $n+1$ is even, then n is odd [$\to$I].

The step "n even, so $n+1$ odd" plays the part that (E3) played in the theorem: it is the implication fired by $\to$E inside the impossible case. Here the impossible case is the **first** one, and the other case is the goal.

Exercise. B3: Two vacuous promises

Let x be a real number. Prove that if $x^2 + 1 \leq 0$, then $x > 100$; and that if $x^2 + 1 \leq 0$, then

$x < -100$. Write each as a derivation with one premise, a school fact, and explain why the two statements together are no paradox.

Solution. The school fact is $x^2+1 > 0$, true for every real x since $x^2 \geq 0$ gives $x^2+1 \geq 1 > 0$. For real numbers, $x^2 + 1 > 0$ says the same as $\neg(x^2 + 1 \leq 0)$. The first derivation is the pattern of Example (An implication with an impossible hypothesis):

1	$\neg(x^2 + 1 \leq 0)$	
2	$\quad x^2 + 1 \leq 0$	asm, for $\to$I
3	$\quad \bot$	$\neg$E 2, 1
4	$\quad x > 100$	$\bot$E 3
5	$x^2 + 1 \leq 0 \to x > 100$	$\to$I 2–4

The second is the same, with $x < -100$ on lines 4 and 5: $\bot$E may write any statement.

No real number is both greater than 100 and less than -100. That is no paradox, because neither implication ever applies: its hypothesis holds for no real x. Both are vacuously true, and both say nothing about x, as the warning on vacuous truth says.

Exercise. B4: An impossible case, scaffolded

A search for $P \vee Q,\ P \to R,\ Q \to S,\ \neg R \vdash S$ has reached the stage below. *Name the move, then complete it.*

1	$P \vee Q$	
2	$P \to R$	
3	$Q \to S$	
4	$\neg R$	
5	$\quad P$	asm, for $\vee$E
	$\quad \vdots$	
8	$\quad S$	?
9	$\quad Q$	asm, for $\vee$E
	$\quad \vdots$	
10	$\quad S$	?
11	S	$\vee$E 1, 5–8, 9–10

Solution. The move is *an impossible case closes itself.* The case P cannot reach S honestly: nothing in it produces S. So it reaches a clash: P fires line 2, and R clashes with line 4. The case Q reaches S by firing line 3.

1	$P \lor Q$	
2	$P \to R$	
3	$Q \to S$	
4	$\neg R$	
5	$\quad P$	asm, for $\lor$E
6	$\quad R$	$\to$E 2, 5
7	$\quad \bot$	$\neg$E 6, 4
8	$\quad S$	$\bot$E 7
9	$\quad Q$	asm, for $\lor$E
10	$\quad S$	$\to$E 3, 9
11	S	$\lor$E 1, 5–8, 9–10

Exercise. B5: Derive or break

For each claim, either derive it with the rules of Chapter 1 and of this section ($\neg$E and $\bot$E; no other rule for "not" yet), or break it by statements about a real number x that make its counterexample set true at one value of x.

(a) $(P \land \neg Q) \lor R,\ Q \vdash R$.
(b) $P \lor Q,\ \neg Q \vdash \neg P$.
(c) $P \to \neg P,\ P \vdash Q$.
(d) $P \to Q,\ \neg P \vdash \neg Q$.
(e) $\neg(P \land Q),\ P \vdash Q \to R$.
(f) $\neg(P \land Q) \vdash \neg P \lor Q$.
(g) $P \lor Q,\ P \to R,\ \neg R \vdash P$.
(h) $P \to (Q \land \neg R),\ R \vdash P \to S$.

Solution.

(a) Derive. Split on the first premise; the case $P \land \neg Q$ is impossible beside the premise Q, and the case R is the goal.

1	$(P \wedge \neg Q) \vee R$	
2	Q	
3	$\quad P \wedge \neg Q$	asm, for $\vee$E
4	$\quad \neg Q$	$\wedge$E 3
5	$\quad \bot$	$\neg$E 2, 4
6	$\quad R$	$\bot$E 5
7	$\quad R$	asm, for $\vee$E
8	R	$\vee$E 1, 3–6, 7–7

(b) Break. Take P: "$x > 0$", Q: "$x > 5$". At $x = 1$, $P \vee Q$ is true ($1 > 0$), $\neg Q$ is true ($1 > 5$ fails), and P is true, so $\neg\neg P$ is true: the counterexample set $P \vee Q$, $\neg Q$, $\neg\neg P$ is true.

(c) Derive. The premise P fires the implication, and its conclusion clashes with P.

1	$P \to \neg P$	
2	P	
3	$\neg P$	$\to$E 1, 2
4	$\bot$	$\neg$E 2, 3
5	Q	$\bot$E 4

(d) Break. Take P: "$x > 3$", Q: "$x > 1$". The premise $P \to Q$ holds for every x. At $x = 2$, $\neg P$ is true ($2 > 3$ fails) and Q is true ($2 > 1$), so $\neg\neg Q$ is true: the counterexample set $P \to Q$, $\neg P$, $\neg\neg Q$ is true.

(e) Derive. The goal is an implication; inside its box, $\wedge$I manufactures the other half of the clash with the first premise.

1	$\neg(P \wedge Q)$	
2	P	
3	$\quad Q$	asm, for $\to$I
4	$\quad P \wedge Q$	$\wedge$I 2, 3
5	$\quad \bot$	$\neg$E 4, 1
6	$\quad R$	$\bot$E 5
7	$Q \to R$	$\to$I 3–6

(f) Break. Take P: "$x > 0$", Q: "$x > 5$". At $x = 1$, $P \wedge Q$ is false, so $\neg(P \wedge Q)$ is true; and $\neg P$ and Q are both false, so $\neg P \vee Q$ is false. The counterexample set $\neg(P \wedge Q)$, $\neg(\neg P \vee Q)$ is true.

(g) Break. Take P: "$x > 5$", Q: "$x > 0$", R: "$x > 3$". The premise $P \to R$ holds for every x. At $x = 1$, $P \vee Q$ is true ($1 > 0$), $\neg R$ is true ($1 > 3$ fails), and P is false, so $\neg P$ is true: the counterexample set $P \vee Q$, $P \to R$, $\neg R$, $\neg P$ is true.

(h) Derive. Inside the box of P, the first premise fires, and its half $\neg R$ clashes with the premise R.

1	$P \to (Q \wedge \neg R)$	
2	R	
3	$\quad P$	asm, for $\to$I
4	$\quad Q \wedge \neg R$	$\to$E 1, 3
5	$\quad \neg R$	$\wedge$E 4
6	$\quad \bot$	$\neg$E 2, 5
7	$\quad S$	$\bot$E 6
8	$P \to S$	$\to$I 3–7

The four that break are look-alikes. (b) draws the wrong conclusion from a refuted disjunct: the disjunctive syllogism of Section 2.3, after putting the disjunction in the other order, gives P, not $\neg P$. (d) denies the antecedent, the look-alike of the modus tollens that Section 2.3 derives. (g) is a look-alike of Example (An impossible case), whose conclusion is the other disjunct. (f) is a look-alike of a De Morgan law of Section 2.5. The four that derive follow worked patterns of this section: Example (An impossible case) for (a), Proposition 2.1.1 for (c), Proposition 2.1.2 for (e) and (h).

Marking scheme. (16 points) For each claim: 1 for the verdict; 1 for a derivation every line of which checks, or for a choice of statements and one value of x at which **every** member of the counterexample set is true. Fatal errors (0 for the claim): a "break" at which some premise is false; a derivation that uses a rule not yet available ($\neg$I, RAA or a derived rule of Section 2.3).

Exercise. B6: Name the rule you reject

(a) List the rules used by a derivation of $\neg P \vdash P \to Q$ and by one of $Q \vdash P \to Q$.

(b) Suppose $\bot$E were dropped. Derive $\neg P \vdash P \to \bot$ without it, and say what this derivation stops short of.

(c) Suppose instead $\to$I were allowed only when the box's assumption is cited by some line of the box or is itself the box's last line. Which of the two derivations of (a) is lost? Is the one-line box of $\vdash P \to P$ lost? Is the sequent $Q \vdash P \to Q$ lost?

(d) Let n be an integer. For $n = 3$, $n = 6$ and $n = 8$, write the instance of "if n is a multiple of 4, then n is even", and say for each whether the first sequent of (a), the second, or neither, decides its truth.

Solution.

(a) $\neg P \vdash P \to Q$: $\to$I, $\neg$E, $\bot$E (Proposition 2.1.2). $Q \vdash P \to Q$: $\to$I and repetition.

(b) Four lines:

1	$\neg P$	
2	$\quad P$	asm, for $\to$I
3	$\quad \bot$	$\neg$E 2, 1
4	$P \to \bot$	$\to$I 2–3

It stops at $\bot$ inside the box, where the derivation of Proposition 2.1.2 wrote Q: the step from $\bot$ to Q was $\bot$E's. What survives says only that P leads to a contradiction.

(c) The derivation of $Q \vdash P \to Q$ is lost: no line of its box cites the assumption P (line 3 repeats line 1). The box of $\vdash P \to P$ is kept: its last line is its assumption, which the restriction allows. The derivation of Proposition 2.1.2 is kept: line 3 cites the assumption. But the sequent $Q \vdash P \to Q$ is not lost. This derivation meets the restriction, since line 3 cites the assumption:

1	Q	
2	$\quad P$	asm, for $\to$I
3	$\quad P \wedge Q$	$\wedge$I 2, 1
4	$\quad Q$	$\wedge$E 3
5	$P \to Q$	$\to$I 2–4

The restriction kills one derivation, not the sequent; a logic that really rejects the paradox has to restrict more than one rule.

(d) $n = 3$: "if 3 is a multiple of 4, then 3 is even"; hypothesis and conclusion are both false, so only $\neg P \vdash P \to Q$ decides it. $n = 6$: "if 6 is a multiple of 4, then 6 is even"; the hypothesis is false and the conclusion true, so either sequent decides it. $n = 8$: "if 8 is a multiple of 4, then 8 is even"; both parts are true ($8 = 4 \cdot 2$), so the first sequent does not apply and the second does, since the conclusion is true. This is the one instance of the three whose hypothesis holds, the kind in which the theorem says something: here it can be fired by $\to$E.

Marking scheme. (10 points) 2: (a), both rule lists. 2: (b), a correct four-line derivation (1) and the remark that the step to Q is $\bot$E's (1). 3: (c), the right derivation lost and the reason, that the assumption is never cited (1); $\vdash P \to P$ kept (1); and a derivation of $Q \vdash P \to Q$ that meets the restriction (1). 3: (d), one point per value: the truth values of hypothesis and conclusion, and the right sequent ($n = 3$: the first only; $n = 6$: either; $n = 8$: the second). Fatal error (0 for (b)): a derivation of $P \to Q$ itself without $\bot$E that cites a line inside a closed box.

2.1.8.3 C. Going deeper

Exercise. C1: A disjunct that cannot hold

(a) Show that $P \vee (Q \wedge \neg Q) \vdash P$.

(b) What does (a) say about the statement "$x > 0$, or $x = 1$ and $x \neq 1$", for a real number x?

Solution.

(a) The premise is a disjunction, so split it by $\vee$E, with both boxes ending in the goal P.

The box of P is a one-line box. The box of $Q \wedge \neg Q$ is impossible: $\wedge$E gives a clash, and $\bot$E closes the case.

1	$P \vee (Q \wedge \neg Q)$	
2	$\quad P$	asm, for $\vee$E
3	$\quad Q \wedge \neg Q$	asm, for $\vee$E
4	$\quad Q$	$\wedge$E 3
5	$\quad \neg Q$	$\wedge$E 3
6	$\quad \bot$	$\neg$E 4, 5
7	$\quad P$	$\bot$E 6
8	P	$\vee$E 1, 2–2, 3–7

(b) With P: "$x > 0$" and Q: "$x = 1$", the statement is $P \vee (Q \wedge \neg Q)$, since $x \neq 1$ is $\neg(x = 1)$. By (a), it gives $x > 0$. So the second disjunct adds nothing: a disjunct that can never hold is an impossible case, and it closes itself. (Conversely, $x > 0$ gives the whole statement by $\vee$I, so the two say the same thing.)

Exercise. C2: A clash that escaped its box

A student wants to show $\neg P \vdash P \to Q$ and writes:

1	$\neg P$	
2	$\quad P$	
3	$\quad \bot$	$\neg$E 2, 1
4	Q	$\bot$E 3
5	$\quad P$	asm, for $\to$I
6	$\quad Q$	R 4
7	$P \to Q$	$\to$I 5–6

(a) Is the sequent $\neg P \vdash P \to Q$ derivable?

(b) Find the first line that is not justified, and name the clause of Definition 1.4.2 it breaks.

(c) Show that the move made at that line can lead from a true premise to a false conclusion.

(d) Write a correct derivation.

Solution.

(a) Yes: it is Proposition 2.1.2. The conclusion is true; the question is whether this argument proves it.

(b) Line 4. It cites line 3, which lies inside the box 2–3, and that box has closed before line 4. By the scope rule of Definition 1.4.2, a line inside a closed box is not available. (Lines 1–3 are fine: a box may close without a rule discharging it, although nothing then leaves it. Lines 5–7 would be fine if line 4 were.)

(c) Lines 1–4 alone would "derive" $\neg P \vdash Q$ by the same move. Read P as "$x > 0$" and Q as "$x = 7$", and take $x = -1$. The premise $\neg(x > 0)$ is true, and the conclusion $x = 7$ is false. So the move "write, outside a box, what a closed box's $\bot$ would allow" can lead from a true premise to a false line. The line $\bot$ on line 3 holds only under the assumption P; once the box closes, that assumption is gone, and so is the clash.

(d) Move the $\bot$E inside the box, where the clash is:

1	$\neg P$	
2	$\quad P$	asm, for $\to$I
3	$\quad \bot$	$\neg$E 2, 1
4	$\quad Q$	$\bot$E 3
5	$P \to Q$	$\to$I 2–4

Marking scheme. (10 points) - 2: answers (a) yes, citing Proposition 2.1.2 or giving a derivation. - 3: names line 4 as the first failure and the scope rule (a line inside a closed box is not available) as the clause broken. - 3: gives a counterexample **to the step**, not only to some conclusion: a reading of P, Q under which lines 1–4 go from a true premise to a false line. - 2: a correct derivation, with every citation available.

Fatal errors (0 for the step): blaming line 7 or line 6 (they are correct once line 4 is); saying that $\bot$E itself is the error.

2.2 Proving a Negation

The rules of Section 2.1 **use** a negation: $\neg$E records a clash, and $\bot$E turns a clash into any statement we like. Neither rule can **prove** a negation unless the lines already clash: $\bot$E writes one only from a line $\bot$. So we still cannot prove that no rational number has square 2, or even that $x^2 + 1$ is never 0. This section adds the one rule that proves a negation, $\neg$I, and uses it for one of the oldest proofs in mathematics.

2.2.1 How to prove that something is not so

To show that no rational number has square 2, we cannot inspect the rational numbers one by one: there are infinitely many. A mathematician instead supposes the opposite and shows that it leads to something impossible: "suppose r is rational and $r^2 = 2$; ...; this is a contradiction; hence $r^2 \neq 2$."

Try the smaller goal $x^2 + 1 \neq 0$, for a real number x, with the rules so far. The goal is $\neg(x^2 + 1 = 0)$, and its main connective is $\neg$. Chapter 1's goal-directed move asks for the introduction rule of the main connective, and $\neg$ has none yet. The rules of Section 2.1 do not help: $\bot$E could write the goal, but only from a line $\bot$, and nothing available clashes. The statement $x^2 + 1 = 0$ is not a line of anything; it is the thing we want to rule out.

The mathematician's sentence shows the missing rule. "Suppose $r^2 = 2$" opens a box, exactly like the box of $\to$I in Definition 1.4.1, and "this is a contradiction" is its last line, $\bot$. The only new thing is what the box gives when it closes.

To prove "not φ", assume φ and reach a contradiction.

Definition 2.2.1. Rules for Negation

i	φ	asm, for $\neg$I
	$\vdots$	
j	$\bot$	
k	$\neg\varphi$	$\neg$I i–j

i	φ	
j	$\neg\varphi$	
k	$\bot$	$\neg$E i, j

The left schema is $\neg$I (**not-introduction**): the assumption φ is **discharged** when the box closes, and the line $\bot$ is a line of that box itself, not of a box inside it. The right schema is $\neg$E, recalled from Definition 2.1.1, so that the rule that proves a negation and the rule that uses one stand together.

This rule has conditions that matter; read each one.

1. **Open a box** whose first line is the assumption φ. It is the **unnegated** statement: the thing we want to refute, not the thing we want to prove. Any statement may be assumed. In a finished derivation this assumption is printed **"asm, for $\neg$I"**: the annotation names the rule that will close the box, as "asm, for $\to$I" did in Section 1.4.
2. **Inside the box** every rule may be used, and every line available just above the box may be cited, as in any box (Definition 1.4.2).
3. **The last line of the box** is exactly $\bot$, and it is a line of that box itself, not inside a further box opened within it. This is the same wording as for $\to$I, and every box rule shares it.
4. **Close the box** and write $\neg\varphi$ on the next line, outside it. The justification cites the box as a range, "$\neg$I i–j". The assumption is discharged: after the box closes, nothing inside it may be cited, and only the negation leaves.
5. **The new line is always a negation.** $\neg$I never writes a statement whose main connective is not $\neg$.

↻ **Second reading** · *Edge cases of not-introduction*

The degenerate case is a box of one line, whose assumption is already $\bot$. The box 1–1 then ends on $\bot$ as soon as it opens, and $\neg$I gives $\neg\bot$:

1	$\bot$	asm, for $\neg$I
2	$\neg\bot$	$\neg$I 1–1

So $\vdash \neg\bot$: "it is not the case that contradiction" is derivable from nothing, as it should be.

A noninstance by one change. Take a box opened by φ whose last line is $\neg\varphi$ instead of $\bot$. It looks as if the assumption has refuted itself, and $\neg$I still does not apply: the box must end on $\bot$. One more line, $\bot$ by $\neg$E citing the assumption and the line $\neg\varphi$, repairs it. The same holds if $\bot$ is reached only inside a smaller box within the box: it is not then a line of the box itself.

A box that never uses its assumption. A box may reach $\bot$ without ever using its assumption. That is legal, but it is a sign: if no line of the box depends on φ, the same lines could have been written just above the box, where $\bot$ would then stand, and $\bot$E would have written $\neg\varphi$, or anything else, directly. The lines available above the box were already contradictory. When

your ¬I box never uses its assumption, look for the error above the box, not in it.

Why this rule. The statement $\neg\varphi$ says that φ fails, and the box shows what would happen if φ held: together with the lines above it, it would lead to $\bot$, which never holds. So in any situation where the lines above the box hold, φ cannot hold. That is exactly what $\neg\varphi$ claims.

The rules for ¬ have the same shapes as rules we already know, with $\bot$ in one slot. Put them beside the rules for $\to$, applied to the statement $\varphi \to \bot$:

To ...	a negation $\neg\varphi$	an implication $\varphi \to \bot$
prove it	box from φ to $\bot$, then ¬I	box from φ to $\bot$, then $\to$I
use it, given φ	$\bot$ by ¬E	$\bot$ by $\to$E

Remark. Row by row, $\neg\varphi$ behaves exactly like $\varphi \to \bot$: "not φ" means "φ leads to contradiction". Some presentations of logic take this as the definition of ¬. This book keeps ¬ as a connective with its own rules.

Warning. Do not open the box with what you want to prove. To prove $\neg\varphi$, the box opens with φ. A box opened by $\neg\varphi$ and ending in $\bot$ gives $\neg\neg\varphi$ by ¬I, the negation of the goal, not the goal. So a box opened with $\neg(x > 0)$ and closed by ¬I gives $\neg\neg(x > 0)$, not $x > 0$; Section 2.5 shows the other kind of "suppose not". This is the negation's version of the converse error that Section 1.4 found in Argument B.

In the Givens/Goal table, ¬I moves φ into the Givens and makes $\bot$ the goal.

2.2.2 The first example: a number that is never zero

The goal we could not reach at the start of the section comes first.

Example. A negative statement proved by not-introduction

Let x be a real number. Prove that $x^2 + 1 \neq 0$, naming the rules.

Idea. The goal $x^2 + 1 \neq 0$ is $\neg(x^2 + 1 = 0)$, a negation, so open a box on $x^2 + 1 = 0$ and look for a clash. The school fact $x^2 \geq 0$, that is $\neg(x^2 < 0)$, is the negation available; its unnegated half $x^2 < 0$ is the subgoal, and algebra on the assumption produces it.

Solution. The goal $x^2 + 1 \neq 0$ is $\neg(x^2 + 1 = 0)$, so we use ¬I. Suppose $x^2 + 1 = 0$ [assumption for ¬I]. Then $x^2 = -1$, so $x^2 < 0$. But $x^2 \geq 0$, that is, $\neg(x^2 < 0)$, so we have a contradiction [¬E]. Hence $x^2 + 1 \neq 0$ [¬I].

With the school fact as the premise:

1	$\neg(x^2 < 0)$	
2	$\quad x^2 + 1 = 0$	asm, for $\neg$I
3	$\quad x^2 < 0$	algebra 2
4	$\quad \bot$	$\neg$E 3, 1
5	$\neg(x^2 + 1 = 0)$	$\neg$I 2–4

Quick check. In this derivation, which box does $\neg$I cite, and which statement opened it?

Solution. The box 2–4, opened by $x^2 + 1 = 0$, the unnegated statement. The conclusion on line 5 is its negation, $x^2 + 1 \neq 0$.

2.2.3 First derivations with not-introduction

The goal-directed move of this section: *a goal $\neg\varphi$ opens a box with φ, and the new goal is $\bot$.* From there, Section 2.1's move takes over: *goal $\bot$, find a clash*, among the available lines or by manufacturing its missing half.

Example. Two negations from one premise

Show that (a) $P \vdash \neg\neg P$, and (b) $(P \vee Q) \to \neg P \vdash \neg P$.

Idea.

(a) The goal $\neg\neg P$ is $\neg(\neg P)$, a negation of $\neg P$. So the box opens with what is negated, $\neg P$, and aims at $\bot$. The clash is immediate: the premise P and the assumption $\neg P$. The goal alone (left); then the box opened on $\neg P$, its goal $\bot$ at its foot (right).

1	P
	$\vdots$
4	$\neg\neg P$

1	P	
2	$\quad \neg P$	asm, for $\neg$I
	$\quad \vdots$	
3	$\quad \bot$	
4	$\neg\neg P$	$\neg$I 2–3

(b) The goal $\neg P$ opens a box on P. Inside it the premise is an implication, and the premise-directed move feeds it its hypothesis $P \vee Q$, which $\vee$I manufactures from P. The premise then hands back $\neg P$, and that clashes with the box's own assumption. The assumption is used twice: once to build the hypothesis, once in the clash.

The box opened, its goal $\bot$ at its foot (left); then the hypothesis $P \vee Q$ manufactured, with line 5 still to reach (right).

1	$(P \vee Q) \to \neg P$	
2	$\quad P$	asm, for $\neg$I
	$\quad \vdots$	
5	$\quad \bot$	
6	$\neg P$	$\neg$I 2–5

1	$(P \vee Q) \to \neg P$	
2	$\quad P$	asm, for $\neg$I
3	$\quad P \vee Q$	$\vee$I 2
	$\quad \vdots$	
5	$\quad \bot$	
6	$\neg P$	$\neg$I 2–5

Solution.

(a) Four lines.

1	P	
2	$\quad \neg P$	asm, for $\neg$I
3	$\quad \bot$	$\neg$E 1, 2
4	$\neg\neg P$	$\neg$I 2–3

(b) Six lines.

1	$(P \vee Q) \to \neg P$	
2	$\quad P$	asm, for $\neg$I
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad \neg P$	$\to$E 1, 3
5	$\quad \bot$	$\neg$E 2, 4
6	$\neg P$	$\neg$I 2–5

In (b), line 4 already says $\neg P$, but it lies inside the box of P and depends on that assumption, so it cannot be the conclusion. Line 6 is the same statement outside the box, where it depends on the premise alone.

Derivation (a) adds two negations. The reverse passage, from $\neg\neg P$ alone to P, would have to **remove** two, and no rule so far has a negation as its one given line and the unnegated statement as its new line: $\neg$E needs a second line and writes only $\bot$, and $\neg$I only adds a $\neg$. Section 2.5 returns to this.

Quick check. Is $\neg(P \vee Q) \vdash \neg P$? If so, give a derivation and name its rules. Compare it with Example (Manufacturing the other half of a clash), which has the same premise.

Solution. Yes. The goal $\neg P$ is a negation, so assume P and aim at $\bot$; $\vee$I turns P into $P \vee Q$, which clashes with the premise.

1	$\neg(P \vee Q)$	
2	$\quad P$	asm, for $\neg$I
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad \bot$	$\neg$E 3, 1
5	$\neg P$	$\neg$I 2–4

The rules are $\vee$I, $\neg$E, and $\neg$I. The box is the same as in Example (Manufacturing the other half of a clash) up to its line $\bot$. There the goal was $P \to R$, so the box had to end on R, written by $\bot$E, and closed by $\to$I. Here the goal is $\neg P$, so the box ends on $\bot$ itself and closes by $\neg$I. In words: if neither P nor Q holds, then in particular P does not.

In the next derivation the clash itself creates a goal that is a negation, and the move of this section is used a second time, one box inside another.

Example. Three negations become one

Show that $\neg\neg\neg P \vdash \neg P$.

Idea. The goal $\neg P$ opens a box on P, with goal $\bot$. Which clash? The only negation available is the premise $\neg(\neg\neg P)$, so the premise-directed question asks for its unnegated half, $\neg\neg P$: that becomes the subgoal. It is itself a negation, $\neg(\neg P)$, so the goal-directed move opens a second $\neg$I box, on $\neg P$, with goal $\bot$. Inside it, $\neg P$ clashes with the outer assumption P. This is the first derivation in which a $\neg$I box is opened for a subgoal that a clash created.

The outer box opened, its goal $\bot$ at its foot (left); then the clash with the premise chosen, so that $\neg\neg P$ becomes the goal of the gap (right).

1	$\neg\neg\neg P$	
2	$\quad P$	asm, for $\neg$I
	$\quad \vdots$	
6	$\quad \bot$	
7	$\neg P$	$\neg$I 2–6

1	$\neg\neg\neg P$	
2	$\quad P$	asm, for $\neg$I
	$\quad \vdots$	
5	$\quad \neg\neg P$	
6	$\quad \bot$	$\neg$E 5, 1
7	$\neg P$	$\neg$I 2–6

Solution. Seven lines.

1	$\neg\neg\neg P$	
2	$\quad P$	asm, for $\neg$I
3	$\quad\quad \neg P$	asm, for $\neg$I
4	$\quad\quad \bot$	$\neg$E 2, 3
5	$\quad \neg\neg P$	$\neg$I 3–4
6	$\quad \bot$	$\neg$E 5, 1
7	$\neg P$	$\neg$I 2–6

Line 4 may cite line 2, because the only box containing line 2, the box 2–6, also contains line 4 (Definition 1.4.2). The inner box 3–4 ends on $\bot$ as a line of its own, and so does the outer box 2–6, at line 6; each $\neg$I cites its own box. Notice what did **not** happen: no line removes a pair of negations. Three became one by adding them, twice, and clashing.

2.2.4 How the rule looks in a mathematical proof

The proof that no rational number has square 2 has the shape of $\neg$I: the claim is negative, so we assume its opposite and look for $\bot$. The tool that finds the clash is Theorem 2.1.3, and the school fact it needs is about writing fractions:

(L) Every rational number can be written as p/q with p and q integers, $q \neq 0$, and p and q **not both even**.

This holds because a fraction whose numerator and denominator are both even can be reduced by 2, and the reduction cannot go on for ever, since the denominator shrinks in absolute value each time. Making "cannot go on for ever" precise needs a fact about the natural numbers that this book does not prove; here we take (L) as known.

Theorem 2.2.1. Irrationality of the Square Root of Two

There is no rational number whose square is 2. So if a real number $\sqrt{2}$ with square 2 exists, it is irrational.

Idea. Fix a rational number r and prove the negative statement $r^2 \neq 2$ by $\neg$I: assume $r^2 = 2$ and hunt for $\bot$. Write $r = p/q$ as in (L), so that the equation becomes $p^2 = 2q^2$. The only available negation is "not both even", so the clash must come from showing that p and q **are** both even; Theorem 2.1.3 is the tool for that, used twice. *Before reading on: the second use of Theorem 2.1.3 needs the line "q^2 is even". From which two earlier lines of the proof will it come?*

Proof. Let r be a rational number, and suppose $r^2 = 2$ [assumption for $\neg$I]. By (L), $r = p/q$ for integers p and q with $q \neq 0$, not both even. Then $p^2 = r^2q^2 = 2q^2$, so p^2 is even. By Theorem 2.1.3, p is even [$\to$E], say $p = 2m$ for an integer m. Then $4m^2 = 2q^2$, so $q^2 = 2m^2$, and q^2 is even. By Theorem 2.1.3 again, q is even [$\to$E]. So p and q are both even [$\wedge$I], while they were chosen not both even; this is a contradiction [$\neg$E]. Hence $r^2 \neq 2$ [$\neg$I].

Since r was an arbitrary rational number, no rational number has square 2. This proves the theorem.

■

Quick check. In the proof, which step is ¬E, and which two statements does it cite?

Solution. The step "this is a contradiction". It cites "p and q are both even", derived inside the box of $r^2 = 2$ by ∧I, and "p and q are not both even", which (L) supplied when p and q were chosen. In the derivation below these are lines 11 and 2, and the step is line 12.

The propositional part of the proof fits in a derivation. Fix r and the integers p, q that (L) supplies. Lines 1 and 2 record what (L) says about them; lines 3 and 4 are Theorem 2.1.3 for $n = p$ and $n = q$; the lines marked "algebra" are the computations of the proof.

1	$r = p/q,\ q \neq 0$	
2	$\neg(p \text{ even} \wedge q \text{ even})$	
3	p^2 even $\to p$ even	
4	q^2 even $\to q$ even	
5	$r^2 = 2$	asm, for ¬I
6	$p^2 = 2q^2$	algebra 1, 5
7	p^2 even	def. of even 6
8	p even	→E 3, 7
9	q^2 even	algebra 6, 8
10	q even	→E 4, 9
11	p even ∧ q even	∧I 8, 10
12	$\bot$	¬E 11, 2
13	$\neg(r^2 = 2)$	¬I 5–12

Line 13 is $r^2 \neq 2$. The box 5–12 is a ¬I box as the schema asks: opened by the unnegated statement $r^2 = 2$, and ending on $\bot$ as a line of its own. Three steps of the proof lie outside the derivation: applying Theorem 2.1.3, a fact about every integer, to p and to q; choosing p, q and m; and passing from the arbitrary r to every rational number. All three are steps about "some" and "every", used as ordinary reasoning (the note at the start of this chapter).

Notice also what the proof does **not** need: that $\sqrt{2}$ exists as a real number. The theorem is a statement about rational numbers only.

And notice which rule closes it. "Suppose $r^2 = 2$" opens a box whose **negation** is the goal, so the proof is ¬I, even though it is what everyone calls a proof by contradiction.

2.2.5 The rule card

Before the card, the rule at work on large statements. Each row is one use of ¬I: the box, and the line it writes.

Example. Instances of ¬I

φ	The box	New line
$P \vee \neg Q$	from $P \vee \neg Q$ (asm) to $\bot$	$\neg(P \vee \neg Q)$
$Q \to R$	from $Q \to R$ (asm) to $\bot$	$\neg(Q \to R)$
$\neg(P \wedge R)$	from $\neg(P \wedge R)$ (asm) to $\bot$	$\neg\neg(P \wedge R)$

Not an instance. From the third box, the line "$P \wedge R$ — ¬I" is not an instance: ¬I puts one ¬ in front of the whole assumption and never takes one away. (The rule that writes $P \wedge R$ from that box is Section 2.5's.)

Chapter 1's seven rules are on the card at the end of Section 1.6; here are this chapter's so far.

Rule	Shape	You write	Lean
¬E	φ, $\neg\varphi \Longrightarrow \bot$	"…, which contradicts …"	`exact hn hp` (goal `False`)
⊥E	$\bot \Longrightarrow$ any φ	"This case cannot occur." "So there is nothing to prove."	`exfalso`
¬I	box φ … $\bot \Longrightarrow \neg\varphi$	"Suppose P. … This is a contradiction. Hence not P."	`intro hp` (goal `False`)

In Lean, the negation of `P` is defined to be `P → False`, with `False` playing the part of ⊥: the reading of the remark above. So `intro` performs both →I and ¬I.

Warning (Not allowed: ¬I). Each step below is wrong; the tag names the clause of the rule's reading that it breaks.

- Close the box on $\neg\varphi$ instead of $\bot$: clause 3, the last line is $\bot$ (one ¬E line repairs it; the fold "Edge cases of not-introduction").
- Close it on a $\bot$ that stands in a box inside it: clause 3, the last line is a line of the box itself.
- Write φ after a box opened by $\neg\varphi$: clause 4, the new line is the negation of the whole assumption, here $\neg\neg\varphi$.
- Write the negation of part of the assumption: clause 4, the whole assumption is negated (the pair below).
- Cite a line of the box after it has closed: clause 4, only the negation leaves.

The discharge is where the rule bites, so here is one wrong step beside the right one, with the premise $\neg(Q \vee R)$. On the left, line 6 negates only P, part of the assumption; on the right, line 6 negates the whole assumption.

1	$\neg(Q \vee R)$	
2	$\quad P \wedge Q$	asm, for $\neg$I
3	$\quad Q$	$\wedge$E 2
4	$\quad Q \vee R$	$\vee$I 3
5	$\quad \bot$	$\neg$E 4, 1
6	$\neg P$	$\neg$I 2–5

1	$\neg(Q \vee R)$	
2	$\quad P \wedge Q$	asm, for $\neg$I
3	$\quad Q$	$\wedge$E 2
4	$\quad Q \vee R$	$\vee$I 3
5	$\quad \bot$	$\neg$E 4, 1
6	$\neg(P \wedge Q)$	$\neg$I 2–5

The left list would show $\neg(Q \vee R) \vdash \neg P$, which breaks. Let x be a real number, with P: "$x > 0$", Q: "$x > 5$", R: "$x < -5$". At $x = 1$ the premise is true and $\neg P$ is false.

Remark (Looking ahead). *Sets, Relations, and Cardinality*, Chapter 2 proves (L) from a fact about the natural numbers called well-ordering. *Analysis on the Real Line*, Chapter 1 uses this section's theorem in exactly the form proved here, as evidence that the rational numbers have a gap where $\sqrt{2}$ should be.

2.2.6 Exercises

2.2.6.1 A. Check your understanding

Exercise. A1

(a) Write the schema of $\neg$I from memory, and say which statement opens its box.

(b) Which rule does this sentence use, and which box does it close? "Suppose 0 had a reciprocal y. Then $1 = 0 \cdot y = 0$, which is absurd. So 0 has no reciprocal."

(c) In the proof of Theorem 2.2.1, the sentence "p is even" follows from "p^2 is even". Which result and which rule justify it?

(d) Where in the proof of Theorem 2.2.1 is (L) used, and what would go wrong if (L) said only "$r = p/q$ with $q \neq 0$"?

(e) Run the proof of Theorem 2.2.1 with 4 in place of 2: assume $r^2 = 4$ and write $r = p/q$ as in (L). Which step can no longer be justified? (The conclusion "no rational number has square 4" is false, so some step must fail.)

Solution.

(a) The schema is the left one of Definition 2.2.1: a box opened by the assumption φ, ending on $\bot$ as a line of the box itself, then $\neg\varphi$ outside it, citing the box. The box opens with the **unnegated** statement φ.

(b) $\neg$I. The conclusion is the negative statement "0 has no reciprocal"; the box opens with "0 has a reciprocal y", and "which is absurd" is its line $\bot$, reached by $\neg$E from $1 = 0$ and the fact $1 \neq 0$.

(c) Theorem 2.1.3, for $n = p$, is the implication "if p^2 is even, then p is even"; together with "p^2 is even" it gives "p is even" by $\to$E (line 8 of the derivation, citing lines 3 and 7).

(d) The fact (L) supplies the line "p and q are not both even" (line 2 of the derivation), which the ¬E of line 12 cites. Without that clause the proof still derives "both even", but it contradicts nothing ($2 = 4/2$ has both even), and no $\bot$ is reached.

(e) With $r^2 = 4$ we get $p^2 = 4q^2 = 2(2q^2)$, so p^2 is even and p is even, $p = 2k$. Then $4k^2 = 4q^2$, so $q^2 = k^2$. This is where the proof stalls: nothing says q^2 is even, so Theorem 2.1.3 cannot be applied to q. Indeed for $r = 2 = 2/1$ we have $q = 1$, which is odd. Every step before the stall is correct; the proof simply fails to reach $\bot$.

Exercise. A2: Drill

The parts run from one step, through a missing line, a justification column and steps to judge, to short derivations.

(a) Which statement opens the ¬I box for the goal $\neg(P \vee Q)$?

(b) A box opens with $\neg P$ and ends with $\bot$. What does ¬I write after it?

(c) For a real number x, the goal is $x \neq 3$. Which statement opens the box?

(d) A box runs from line 4, $(P \to Q) \wedge \neg R$, to line 8, $\bot$, a line of that box itself. What does "¬I 4–8" write?

(e) Line 10 is $\neg(Q \vee \neg P)$, justified "¬I 6–9". What are lines 6 and 9?

(f) Inside a box opened at line 3 by P, line 5 is $\bot$, justified "¬E 4, 2", and line 2 is $\neg(P \vee R)$. What is line 4?

(g) Fill in (i)–(iv), and say what the derivation shows, written with $\vdash$.

1	$\neg Q$	
2	$P \to (Q \wedge R)$	
3	$\quad P$	asm, for ¬I
4	$\quad Q \wedge R$	(i)
5	$\quad Q$	(ii)
6	$\quad \bot$	(iii)
7	$\neg P$	(iv)

(h) A box opens with P and its last line is $\neg P$. May ¬I close it?

(i) Line 1 is the premise $\neg Q$; line 2, Q, opens a box; line 3, $\bot$, justified "¬E 2, 1", is the last line of that box. What may line 4, just after the box, be by ¬I?

(j) Find the illegal line, and say why it is illegal.

1	$P \to Q$	
2	$\neg Q$	
3	$\quad P$	asm, for ¬I
4	$\quad Q$	→E 1, 3
5	$\quad \bot$	¬E 4, 2
6	$\neg P$	¬I 3–5
7	Q	→E 1, 3

(k) Lines 3–7 form a box opened by $P \wedge Q$. Inside it, lines 4–6 form a box opened by R, whose last line, 6, is $\bot$; line 7 is $\neg R$, justified "¬I 4–6". Is "8. $\neg(P \wedge Q)$ — ¬I 3–7" correct?

(l) "Suppose $x^2 + 4 = 2x$ for a real x. Then $(x-1)^2 = -3 < 0$, which contradicts $(x-1)^2 \geq 0$." What may "Hence" write, and by which rule?

(m) Derive $\neg P \vdash \neg(P \wedge Q)$.

(n) Derive $P \to \neg P \vdash \neg P$.

(o) Derive $P \wedge \neg Q \vdash \neg(P \to Q)$.

Solution.

(a) $P \vee Q$, the statement the goal negates.

(b) $\neg\neg P$: one ¬ in front of the whole assumption.

(c) $x = 3$, since $x \neq 3$ is $\neg(x = 3)$.

(d) $\neg((P \to Q) \wedge \neg R)$, the negation of the whole assumption of line 4.

(e) Line 6 is the assumption $Q \vee \neg P$, which opens the box; line 9 is $\bot$, the last line of that box.

(f) $P \vee R$, the unnegated half of line 2; it comes from line 3 by ∨I.

(g) Line 3 is the assumption for ¬I; (i) →E 2, 3; (ii) ∧E 4; (iii) ¬E 5, 1; (iv) ¬I 3–6. The derivation shows $\neg Q,\ P \to (Q \wedge R) \vdash \neg P$.

(h) No: the last line must be $\bot$. First write $\bot$ by ¬E, citing the assumption and the line $\neg P$; then ¬I closes the box.

(i) $\neg Q$, the negation of the box's assumption; not Q.

(j) Line 7: it cites line 3, which lies inside the box 3–5, closed before line 7.

(k) No. The box 3–7 ends on $\neg R$, not on $\bot$; the $\bot$ at line 6 is a line of the inner box, not of the box 3–7 itself (clause 3). The outer box needs a line $\bot$ of its own.

(l) $x^2 + 4 \neq 2x$, by ¬I. The supposition opens the box, and the contradiction is its line $\bot$; the step to $(x-1)^2 = -3$ is the identity $x^2 + 4 - 2x = (x-1)^2 + 3$.

(m) 2 $P \wedge Q$, assumption, for ¬I; 3 P, ∧E 2; 4 $\bot$, ¬E 3, 1; 5 $\neg(P \wedge Q)$, ¬I 2–4.

(n) 2 P, assumption, for ¬I; 3 $\neg P$, →E 1, 2; 4 $\bot$, ¬E 2, 3; 5 $\neg P$, ¬I 2–4.

(o) 2 $P \to Q$, assumption, for ¬I; 3 P, ∧E 1; 4 Q, →E 2, 3; 5 $\neg Q$, ∧E 1; 6 $\bot$, ¬E 4, 5; 7 $\neg(P \to Q)$, ¬I 2–6.

2.2.6.2 B. Practice

Exercise. B1: A box inside a box

A search for $P \to \neg Q \vdash Q \to \neg P$ has reached the stage below. *Name the move that opens each box, then complete it.*

1	$P \to \neg Q$	
	⋮	
7	$Q \to \neg P$	?

Solution. The goal is an implication, so the outer box opens on Q, with goal $\neg P$. That goal is a negation, so the inner box opens on P, with goal $\bot$. Inside it, the premise fires on P and gives $\neg Q$, which clashes with the outer assumption Q.

1	$P \to \neg Q$	
2	$\quad Q$	asm, for $\to$I
3	$\quad\quad P$	asm, for $\neg$I
4	$\quad\quad \neg Q$	$\to$E 1, 3
5	$\quad\quad \bot$	$\neg$E 2, 4
6	$\quad \neg P$	$\neg$I 3–5
7	$Q \to \neg P$	$\to$I 2–6

Line 5 may cite line 2 because the box 2–6 contains line 5. Both boxes were read off the goal: the outer one by Chapter 1's move for an implication, the inner one by this section's move for a negation. In words: if P rules out Q, then Q rules out P.

Exercise. B2: A quadratic with no zero

Let x be a real number. Prove that $x^2 - x + 1 \neq 0$, naming the rules.

Hint: complete the square.

Solution. The goal is $\neg(x^2 - x + 1 = 0)$, a negation, so we use the rule $\neg$I. Suppose that $x^2 - x + 1 = 0$, as the assumption for $\neg$I. Completing the square,

$$x^2 - x + 1 = \left(x - \tfrac{1}{2}\right)^2 + \tfrac{3}{4},$$

so $(x - \frac{1}{2})^2 = -\frac{3}{4}$, and therefore $(x - \frac{1}{2})^2 < 0$. But the square of a real number is never negative, that is, $\neg((x - \frac{1}{2})^2 < 0)$; this is a contradiction [$\neg$E]. Hence $x^2 - x + 1 \neq 0$ [$\neg$I].

This is the pattern of Example (A negative statement proved by not-introduction), with the square $(x-\frac{1}{2})^2$ in place of x^2. The identity can be checked by expanding: $(x-\frac{1}{2})^2 = x^2-x+\frac{1}{4}$, and $\frac{1}{4} + \frac{3}{4} = 1$.

Exercise. B3: The square root of three, used later

You may use that every integer can be written as $3k$, $3k+1$ or $3k+2$ with k an integer, in one way only, and that every rational number can be written as p/q with integers p, q, $q \neq 0$, and p, q not both divisible by 3.

(a) Prove that if n is an integer and n^2 is divisible by 3, then n is divisible by 3.

(b) Hence prove that no rational number has square 3.

Solution.

(a) Suppose n^2 is divisible by 3 [assumption for $\to$I], so n^2 has the form $3j$ with j an integer. The integer n has the form $3k$, $3k+1$ or $3k+2$, and we argue by cases [$\vee$E]; a three-way split is two uses of $\vee$E, one inside a box of the other.
Case $n = 3k$. Then n is divisible by 3, as required.
Cases $n = 3k+1$ *and* $n = 3k+2$. Then

$$(3k+1)^2 = 9k^2 + 6k + 1 = 3(3k^2 + 2k) + 1,$$
$$(3k+2)^2 = 9k^2 + 12k + 4 = 3(3k^2 + 4k + 1) + 1,$$

so n^2 has the form $3j'+1$ as well as the form $3j$. An integer has only one of the three forms, so this is a contradiction [$\neg$E], and the case cannot occur. By $\bot$E, n is divisible by 3 in this case too.
All cases give that n is divisible by 3 [$\vee$E], which proves (a) [$\to$I]. The two impossible cases were closed exactly as the odd case in the proof of Theorem 2.1.3.

(b) Let r be a rational number and suppose $r^2 = 3$ [assumption for $\neg$I]. Write $r = p/q$ with p, q not both divisible by 3. Then $p^2 = 3q^2$, so 3 divides p^2, and by (a), $p = 3m$ for an integer m. Then $9m^2 = 3q^2$, so $q^2 = 3m^2$, and by (a) again, 3 divides q. So 3 divides both p and q [$\wedge$I]; this contradicts their choice [$\neg$E]. Hence $r^2 \neq 3$ [$\neg$I]. Since r was an arbitrary rational number, no rational number has square 3.

The form of (L) matters: the argument ends at "3 divides p and q", which clashes only with "not both divisible by 3". The form "not both even" of the text would give no clash, since $p = q = 3$ are both divisible by 3 and not both even.

Exercise. B4: A premise that turns on itself

Show that $P \to Q,\ Q \to \neg P \vdash \neg P$. Then say in one sentence what the sequent means.

Solution. The goal $\neg P$ is a negation, so assume P. The two premises, used in turn, turn P into $\neg P$, and the clash gives $\bot$.

1	$P \to Q$	
2	$Q \to \neg P$	
3	$\quad P$	asm, for $\neg$I
4	$\quad Q$	$\to$E 1, 3
5	$\quad \neg P$	$\to$E 2, 4
6	$\quad \bot$	$\neg$E 3, 5
7	$\neg P$	$\neg$I 3–6

In words: if P leads to Q, and Q rules P out, then P refutes itself, and so P fails. Line 5 is the moment of self-refutation: the assumption has produced its own negation. Line 6 is still needed, because the box must end on $\bot$, not on $\neg P$.

Exercise. B5: From $\sqrt{2}$ to $\sqrt{p}$

Let p be a prime number. You may use two facts.

- (L_p) Every rational number can be written as a/b with integers a, b, $b \neq 0$, and a, b not both divisible by p.
- (E_p) For every integer n, if p divides n^2, then p divides n.

(a) Prove that no rational number has square p, naming the rule at each logical step.

(b) Where does your proof use (E_p), and where (L_p)?

(c) Your proof never used that p is prime, only the two facts. Put 6 and 9 in place of p. (E_6) is true (Exercise C1 of Section 2.4 proves it; take it as given here). Does (L_6) follow from (L)? Is (E_9) true? What does the argument then prove about $\sqrt{6}$, and why does it prove nothing about $\sqrt{9}$?

Solution.

(a) Let r be a rational number, and suppose $r^2 = p$ [assumption for ¬I]. By (L_p), $r = a/b$ with $b \neq 0$ and a, b not both divisible by p. Then $a^2 = r^2b^2 = pb^2$, so p divides a^2, and by (E_p) p divides a [→E]; write $a = pm$. Then $p^2m^2 = pb^2$, so $b^2 = pm^2$, p divides b^2, and by (E_p) again p divides b [→E]. So p divides both a and b [∧I], which contradicts their choice [¬E]. Hence $r^2 \neq p$ [¬I]; since r was arbitrary, no rational number has square p.

(b) (E_p) twice, once for a and once for b, in the places where Theorem 2.2.1 used Theorem 2.1.3. (L_p) when a and b are chosen, and again at the clash, which cites "not both divisible by p".

(c) For 6: (L_6) follows from (L) of the text, since integers a, b that are not both even are not both divisible by 6. With (E_6) as given, the argument proves that $\sqrt{6}$ is irrational; Exercise C1 of this section does it by another route, with (L) and Theorem 2.1.3 in place of (E_6). (E_6) itself is not proved here: its proof is Exercise C1 of Section 2.4. For 9: (E_9) is false, since 9 divides 3^2 but not 3. The proof breaks at its first use of (E_9), as it must, since $9 = 3^2$ and $3 = 3/1$ is rational: with $r = 3/1$, 9 divides $a^2 = 9$, but not $a = 3$.

Marking scheme. (10 points)

- 2: the ¬I frame, a fixed rational r, and a/b from (L_p).
- 2: $a^2 = pb^2$ and the first use of (E_p).
- 2: $b^2 = pm^2$ and the second use.
- 1: the clash with (L_p), ¬I, and the passage from r to every rational number.
- 1: (b).
- 2: (c), (L_6) from (L) with the conclusion about $\sqrt{6}$ (1), and 9 with a witness against (E_9) (1).

Fatal errors (0 for the step): using "p divides n^2, so p divides n" for $p = 9$ or $p = 4$; dividing by p before p is known to divide a.

Exercise. B6: Derive or break

For each claim, either derive it with the rules of Chapter 1 and Sections 2.1–2.2, or break it

by statements about a real number x that make its counterexample set true at one value of x.

(a) $P \vdash \neg(\neg P \wedge Q)$.
(b) $\neg(P \wedge Q) \vdash \neg P$.
(c) $(P \wedge Q) \to \neg P,\ Q \vdash \neg P$.
(d) $P \to (Q \to R),\ \neg R \vdash \neg(P \wedge Q)$.
(e) $\neg Q \vdash \neg(P \to Q)$.
(f) $P \to Q,\ P \to \neg Q \vdash \neg P$.
(g) $\neg\neg(P \wedge Q) \vdash \neg\neg P$.
(h) $\neg\neg P \vdash \neg\neg(P \wedge Q)$.

Solution.

(a) Derive. The goal is a negation, so the box opens on $\neg P \wedge Q$, and its first half clashes with the premise.

1	P	
2	$\quad \neg P \wedge Q$	asm, for $\neg$I
3	$\quad \neg P$	$\wedge$E 2
4	$\quad \bot$	$\neg$E 1, 3
5	$\neg(\neg P \wedge Q)$	$\neg$I 2–4

(b) Break. Take P: "$x > 0$", Q: "$x > 5$". At $x = 1$, $P \wedge Q$ is false ($1 > 5$ fails), so $\neg(P \wedge Q)$ is true, and P is true, so $\neg\neg P$ is true. The counterexample set $\neg(P \wedge Q)$, $\neg\neg P$ (Definition 1.1.3) is true.

(c) Derive. On the assumption P, the premise Q completes the antecedent, and the assumption produces its own negation.

1	$(P \wedge Q) \to \neg P$	
2	Q	
3	$\quad P$	asm, for $\neg$I
4	$\quad P \wedge Q$	$\wedge$I 3, 2
5	$\quad \neg P$	$\to$E 1, 4
6	$\quad \bot$	$\neg$E 3, 5
7	$\neg P$	$\neg$I 3–6

(d) Derive. Inside the box of $P \wedge Q$, feed the premise its two hypotheses one at a time; the R it hands back clashes with $\neg R$.

1	$P \to (Q \to R)$	
2	$\neg R$	
3	$\quad P \wedge Q$	asm, for $\neg$I
4	$\quad P$	$\wedge$E 3
5	$\quad Q \to R$	$\to$E 1, 4
6	$\quad Q$	$\wedge$E 3
7	$\quad R$	$\to$E 5, 6
8	$\quad \bot$	$\neg$E 7, 2
9	$\neg(P \wedge Q)$	$\neg$I 3–8

(e) Break. Take P: "$x > 5$", Q: "$x > 3$". The implication $P \to Q$ holds for every x. At $x = 0$, $\neg Q$ is true, and $P \to Q$ is true, so $\neg\neg(P \to Q)$ is true: the counterexample set $\neg Q$, $\neg\neg(P \to Q)$ is true.

(f) Derive. On the assumption P, the two premises give Q and $\neg Q$.

1	$P \to Q$	
2	$P \to \neg Q$	
3	$\quad P$	asm, for $\neg$I
4	$\quad Q$	$\to$E 1, 3
5	$\quad \neg Q$	$\to$E 2, 3
6	$\quad \bot$	$\neg$E 4, 5
7	$\neg P$	$\neg$I 3–6

(g) Derive. The goal $\neg\neg P$ opens a box on $\neg P$. The only way to use the premise is to clash with it, so $\neg(P \wedge Q)$ becomes a subgoal, and a second box proves it, as in Example (Three negations become one).

1	$\neg\neg(P \wedge Q)$	
2	$\quad \neg P$	asm, for $\neg$I
3	$\quad\quad P \wedge Q$	asm, for $\neg$I
4	$\quad\quad P$	$\wedge$E 3
5	$\quad\quad \bot$	$\neg$E 4, 2
6	$\quad \neg(P \wedge Q)$	$\neg$I 3–5
7	$\quad \bot$	$\neg$E 6, 1
8	$\neg\neg P$	$\neg$I 2–7

(h) Break. Take P: "$x > 0$", Q: "$x > 5$". At $x = 1$, P is true, so $\neg\neg P$ is true; $P \wedge Q$ is false, so $\neg(P \wedge Q)$ is true and $\neg\neg(P \wedge Q)$ is false. The counterexample set $\neg\neg P$, $\neg\neg\neg(P \wedge Q)$ is true.

(i) and (h) are the look-alikes, each moving a negation or a conjunct the wrong way: (b) would refute a conjunct from a refuted conjunction, and (h) would add a conjunct under two negations. (g) keeps both negations because $\neg$I can only add them, as the

warning of this section says. (d) unpacks its nested premise inside the ¬I box, one →E per hypothesis, as the two forms of Proposition 1.4.2 suggest.

Marking scheme. (16 points) For each claim: 1 for the verdict; 1 for a derivation every line of which checks, or for a choice of statements and one value of x at which **every** member of the counterexample set is true. Fatal errors (0 for the claim): a "break" at which some premise is false; a derivation that uses a rule other than those of Chapter 1 and Sections 2.1–2.2.

2.2.6.3 C. Going deeper

Exercise. C1: The square root of six

Prove that no rational number has square 6. Use only the facts (E1)–(E3) of Section 2.1, (L), and Theorem 2.1.3, and name the rule at each logical step.

Solution (Hint 1). Follow the proof of Theorem 2.2.1. After writing $p = 2m$, you will need to show that q is even from $3q^2 = 2m^2$, where q^2 is not yet known to be even.

Solution (Hint 2). You will reach $3q^2 = 2m^2$, so $3q^2$ is even. The proof of Theorem 2.1.3 passed from "n^2 is even" to "n is even"; imitate its shape for $3q^2$ and q.

Solution. Let r be a rational number and suppose $r^2 = 6$ [assumption for ¬I]. By (L), $r = p/q$ with integers p, q, $q \neq 0$, not both even. Then $p^2 = 6q^2 = 2(3q^2)$ is even, so p is even by Theorem 2.1.3 [→E]; write $p = 2m$. Then $4m^2 = 6q^2$, that is, $3q^2 = 2m^2$, so $3q^2$ is even.

We show that q is even. By (E1), q is even or odd, and we argue by cases [∨E]. *Case q even:* this is the goal. *Case q odd:* then q^2 is odd by (E3) [→E], say $q^2 = 2j + 1$, and then $3q^2 = 6j + 3 = 2(3j + 1) + 1$ is odd. So $3q^2$ is both even and odd [∧I]; this contradicts (E2) [¬E], so this case cannot occur, and in it q is even [⊥E]. In both cases q is even [∨E].

So p and q are both even [∧I]; this contradicts (L) [¬E]. Hence $r^2 \neq 6$ [¬I], and since r was arbitrary, no rational number has square 6.

Marking scheme. (10 points) - 2: opens with ¬I: a fixed rational r, the assumption $r^2 = 6$, and $r = p/q$ with p, q not both even from (L). - 2: derives p even from $p^2 = 2(3q^2)$ by Theorem 2.1.3. - 3: derives q even from $3q^2 = 2m^2$, either by ruling out the odd case (a correct parity computation and (E2)) or by showing $q^2 = 2(m^2 - q^2)$ even and applying Theorem 2.1.3. - 2: reaches ⊥ by clashing "both even" with (L), and concludes $r^2 \neq 6$ by ¬I. - 1: rules named at the logical steps, and the passage from the arbitrary r to every rational number.

Fatal errors (0 for the step): applying Theorem 2.1.3 to q with only $3q^2$ known to be even; using "if 6 divides n^2, then 6 divides n" without proof.

Exercise. C2: A proof that proves too much

Here is a "proof" that $\sqrt{4}$ is irrational, modeled on Theorem 2.2.1.

> Suppose r is rational and $r^2 = 4$. Write $r = p/q$ with p, q integers, $q \neq 0$, not both divisible by 4. Then $p^2 = 4q^2$, so 4 divides p^2. Since n^2 divisible by 4 implies n divisible by 4, we get $p = 4m$. Then $16m^2 = 4q^2$, so $q^2 = 4m^2$, and as before 4 divides q. This contradicts the choice of p and q. Hence no rational number has square 4.

(a) Explain why the conclusion must be false.

(b) Find the step that fails, and show that it fails with a specific integer.

(c) Every other step is a correct use of a rule. Which rule turns the "contradiction" into the conclusion, and why is the argument still wrong?

Solution.

(a) $2 = 2/1$ is rational and $2^2 = 4$.

(b) The step "n^2 divisible by 4 implies n divisible by 4" is a false lemma: for $n = 2$, $n^2 = 4$ is divisible by 4, but 2 is not. It fails exactly on the witness of (a): with $r = 2/1$ we have $p = 2$, and the lemma would claim that 4 divides 2.

(c) The last step is ¬I, applied to the box opened by "$r^2 = 4$". The rule is applied correctly, but the box contains a line that is not justified: the false lemma was used as if it were a premise available everywhere, through →E. A derivation is only as good as its premises. In Theorem 2.2.1 the corresponding lemma is Theorem 2.1.3, which we proved; here nothing proves the lemma, and nothing could. Compare part (e) of A1, where the honest proof stalls without reaching ⊥; here ⊥ is "reached" through a false line.

2.3 Derived Rules: Proved Sequents as New Rules

Section 2.1 closed an impossible case by hand, and Section 2.2 proved negations by hand, one box at a time. By now the same few patterns keep coming back: the impossible case of Example (An impossible case), and a box that turns a known implication round. Mathematicians give such patterns names ("by contraposition", "since P fails, Q") and use them in one breath. This section proves three of them once, from the rules we have, and then allows each to be used as a single line. The one new thing is that permission: a **derived rule**.

2.3.1 A box that keeps coming back

Here are two short arguments from different corners of mathematics.

> Let n be an integer whose square is not even. Suppose n is even. Then n^2 is even, which contradicts the hypothesis. Hence n is not even.

> Let x be a real number with $x^2 < 1$. Suppose $x \geq 1$. Then $x^2 \geq x \geq 1$, which contradicts $x^2 < 1$. Hence $x < 1$.

The subjects differ; the steps are the same. Each has a known implication ("if n is even, then n^2 is even"; "if $x \geq 1$, then $x^2 \geq 1$"), and a fact saying that the implication's conclusion fails. Each then opens a box on the implication's hypothesis, fires the implication inside it by $\to$E, meets the clash with $\neg$E, and closes the box by $\neg$I (Definition 2.2.1). (For reals, $x^2 < 1$ and $x < 1$ are the exact negations of $x^2 \geq 1$ and $x \geq 1$.) The same four lines appear whenever an implication is used backwards, from the failure of its conclusion to the failure of its hypothesis.

A mathematician does not write these four lines. She writes "since n^2 is not even, n is not even", or "by contraposition", and moves on: the lines are the same every time. But a derivation is checked line by line against the rules, and "by contraposition" is not one of the rules. What licenses the short form? First we derive the pattern once, as a sequent, with the rules we have. Then we say exactly when a sequent already derived may be used as a single line.

2.3.2 Modus tollens, from the rules we have

The pattern of the two arguments is: from an implication and the negation of its conclusion, infer the negation of its hypothesis.

Proposition 2.3.1. Modus Tollens

$P \to Q,\ \neg Q \vdash \neg P$.

Idea. Read the skeleton off the goal. The goal $\neg P$ is a negation, and Section 2.2's move for a negated goal is a box opened on what is negated, here P, with goal $\bot$. Inside the box, look for a clash (Section 2.1's move for the goal $\bot$). The premise $\neg Q$ will clash with Q, so Q is the subgoal, and the premise $P \to Q$ gives it from the box's own assumption. In the pictures, the box on P stands with its goal $\bot$; then $\neg$E writes $\bot$ from the subgoal Q, still unjustified.

1	$P \to Q$	
2	$\neg Q$	
3	$\quad P$	asm, for $\neg$I
	$\quad \vdots$	
5	$\quad \bot$	
6	$\neg P$	$\neg$I 3–5

1	$P \to Q$	
2	$\neg Q$	
3	$\quad P$	asm, for $\neg$I
	$\quad \vdots$	
4	$\quad Q$	
5	$\quad \bot$	$\neg$E 4, 2
6	$\neg P$	$\neg$I 3–5

Proof. The following is a derivation with premises $P \to Q$ and $\neg Q$, and last line $\neg P$.

1	$P \to Q$	
2	$\neg Q$	
3	$\quad P$	asm, for $\neg$I
4	$\quad Q$	$\to$E 1, 3
5	$\quad \bot$	$\neg$E 4, 2
6	$\neg P$	$\neg$I 3–5

Line 5 cites the premise on line 2, which lies in no box and so is available everywhere below it. The box 3–5 ends in $\bot$ as a line of its own, so $\neg$I closes it on line 6. This proves the proposition.

■

The name is Latin for "the mode that takes away": denying the conclusion takes away the hypothesis. Its partner, from $P \to Q$ and P to Q, is $\to$E, traditionally *modus ponens*, "the mode that puts" (Definition 1.3.1). The derivation uses $\to$E, $\neg$E, and $\neg$I, and nothing else.

A second form of the same fact packs the two premises into one implication. It is the form in which mathematicians state it: an implication **implies** its contrapositive.

Proposition 2.3.2. Contrapositive

$P \to Q \vdash \neg Q \to \neg P$.

Idea. The goal is an implication, so Chapter 1's first goal-directed move applies: open a box with its hypothesis $\neg Q$, and the new goal is $\neg P$. That goal is a negation, so open a second box, on P, inside the first. Inside both boxes we are exactly where the proof of Proposition 2.3.1 was: P and $P \to Q$ give Q, which clashes with $\neg Q$.

Proof. The following is a derivation with premise $P \to Q$ and last line $\neg Q \to \neg P$.

1	$P \to Q$	
2	$\quad \neg Q$	asm, for $\to$I
3	$\quad\quad P$	asm, for $\neg$I
4	$\quad\quad Q$	$\to$E 1, 3
5	$\quad\quad \bot$	$\neg$E 4, 2
6	$\quad \neg P$	$\neg$I 3–5
7	$\neg Q \to \neg P$	$\to$I 2–6

Line 5 may cite line 2, because the only box containing line 2, the box 2–6, also contains line 5 (the scope rule of Definition 1.4.2). The inner box closes on line 6, a line of the outer box itself, and line 7 closes the outer box. This proves the proposition.

■

The statement $\neg Q \to \neg P$ is called the **contrapositive** of $P \to Q$. Like modus tollens, its derivation uses only the rules so far: $\to$I, $\to$E, $\neg$I, and $\neg$E. The passage back, from $\neg Q \to \neg P$ to $P \to Q$, is a different matter (Common errors, below).

Now compare the two derivations. Lines 3–6 of the derivation of Proposition 2.3.1 are, line for line and citation for citation, lines 3–6 of the derivation of Proposition 2.3.2: P, then Q by $\to$E 1, 3, then $\bot$ by $\neg$E 4, 2, then $\neg P$ by $\neg$I 3–5. And they are the four lines of the two arguments that opened this section. Writing them out every time buries the argument in bookkeeping.

2.3.3 What a derived rule is

A sequent that has been derived may be used as a rule.

This is what a mathematician does when she writes "by Lemma 3": the lemma was proved once, and each use of it is a single step. For derivations we must say exactly what that step may do.

Definition 2.3.1. Derived Rule

Let $\alpha_1, \ldots, \alpha_n \vdash \beta$ be a sequent, written with statement letters, that has **already been derived** earlier in the book. In a later derivation, replace each letter in $\alpha_1, \ldots, \alpha_n, \beta$ by a statement, the **same** statement for every occurrence of the same letter. If lines $i_1, \ldots, i_n$, all **available** at line k, hold the resulting instances of $\alpha_1, \ldots, \alpha_n$, we may write the resulting instance of β on line k. The justification names the result and cites $i_1, \ldots, i_n$. A sequent used in this way is a **derived rule**. For example, Proposition 2.3.1 becomes the rule MT:

i	$\varphi \to \psi$	
j	$\neg\psi$	
k	$\neg\varphi$	MT i, j

In words: a proved sequent is a new rule. Its premises become the cited lines, its conclusion the line written, and its letters the metavariables of a schema (Definition 1.2.1). Read the definition clause by clause.

1. **"Already derived."** The sequent must have been derived **before** the line that uses it: earlier in the book, and never inside its own derivation. A derivation that used a sequent to derive that same sequent would prove nothing.
2. **"Replace each letter ... the same statement for every occurrence."** The letters of a derived sequent stood for arbitrary statements all along (Definition 1.2.3), so they may be replaced by any statements, provided the replacement is **uniform**. From $(A \wedge B) \to C$ and $\neg C$, the rule MT gives $\neg(A \wedge B)$, with the letters P and Q of Proposition 2.3.1 (written φ and ψ in the schema) replaced by $A \wedge B$ and C.
3. **"Available at line k."** The scope rule of Definition 1.4.2 still governs every citation. A derived rule may not reach into a closed box, just as a primitive rule may not.
4. **"The resulting instance of β."** The rule writes the conclusion, with the same replacement, and nothing else. The derived rule does not open or close any box of the derivation that uses it.

We give the derived rules of this section short tags: MT for Proposition 2.3.1, CP for Proposi-

tion 2.3.2, and, once it is proved below, DS for Proposition 2.3.4. Every other derived sequent is a derived rule too, cited by its name: the four lines 2–5 of Example (An implication with an impossible hypothesis), for instance, are now one line citing Proposition 2.1.2.

Left: with the derived rule. Right: written out.

1	$P \to Q$	
2	$\neg Q$	
3	$\neg P$	MT 1, 2

1	$P \to Q$	
2	$\neg Q$	
3	$\quad P$	asm, for $\neg$I
4	$\quad Q$	$\to$E 1, 3
5	$\quad \bot$	$\neg$E 4, 2
6	$\neg P$	$\neg$I 3–5

Line 3 on the left stands for lines 3–6 on the right: the same premises, the same last line.

A derived rule is shorthand for lines we could always write out, so it lets us prove nothing that the primitive rules could not.

Quick check. In the figure, which lines does line 3 on the left cite, and which lines on the right use those same two premises?

Solution. Lines 1 and 2. On the right, line 4 cites line 1 and line 5 cites line 2.

Second reading · *Why a derived rule adds no power*

Convenience, not power. A derived-rule line abbreviates a stretch of derivation that we could write out in full. To expand it, repeat each cited line by (D3) of Definition 1.4.2, then copy the derivation of the sequent below the repetitions, with its letters replaced, its premises replaced by the repetitions, and its line numbers shifted. Every citation inside the copy points inside the copy or to a repetition, all in the current box, so every citation is available. The copy's boxes close before its last line, which is the instance of β. If the copy itself uses derived rules, expand those lines in the same way; this stops, because each derivation cites only results proved before it. So whatever can be derived with derived rules can be derived with the primitive rules alone.

The degenerate case is a sequent with nothing done to its premise, $P \vdash P$, whose derivation is the single premise line (part (a) of Example (Derivations, simplest first)): as a derived rule it copies an available line, which repetition already does. The first use with content is modus tollens proved a second time, with CP as a rule. The premise $P \to Q$ invites CP, and its conclusion $\neg Q \to \neg P$ is an implication whose hypothesis is the other premise.

1	$P \to Q$	
2	$\neg Q$	
3	$\neg Q \to \neg P$	CP 1
4	$\neg P$	$\to$E 3, 2

Line 3 is the instance of Proposition 2.3.2 with P and Q replaced by themselves; line 4 is modus ponens on it. Written out in full, line 3 would become a repetition of line 1 followed by lines 2–7 of the derivation of Proposition 2.3.2. This is honest because Proposition 2.3.2 was derived from the primitive rules, not from Proposition 2.3.1: the chain of derived rules ends in primitive rules, without a loop.

Warning. A derived rule must match its sequent exactly, under one uniform replacement of letters. From $P \to Q$ and $\neg R$, MT gives nothing: no single statement in place of ψ turns the pair into $\varphi \to \psi$ and $\neg\psi$. Equally, from $P \to \neg Q$ and Q, MT gives nothing yet: it needs the negation of the conclusion, $\neg\neg Q$, and Q is not that line. (Here ¬I would give $\neg\neg Q$ first, as in Example (Two negations from one premise), and then MT applies.) Likewise DS, as stated below, refutes the **left** disjunct; applied blindly to $P \vee Q$ and $\neg Q$, it does not match.

In the Givens/Goal table, a derived rule adds its conclusion to the Givens in one row, as an elimination rule does.

2.3.4 Pasting whole derivations: weakening and cut

The figure expands one derived-rule line inside a derivation. Outside every box there is a cleaner statement, about whole derivations: two derivations can be joined end to end, and premises that are never used do no harm.

Lemma 2.3.3. Weakening and Cut

Let Γ and Δ be lists of statements, and let φ, ψ, $\alpha_1, \ldots, \alpha_m$ $(m \geq 1)$, and β be statements.

(a) **Weakening.** If $\Gamma \vdash \psi$ and every member of Γ is a member of Δ, then $\Delta \vdash \psi$. In particular, $\Gamma \vdash \psi$ implies $\Gamma, \varphi \vdash \psi$.

(b) **Cut.** If $\Gamma \vdash \varphi$ and $\Gamma, \varphi \vdash \psi$, then $\Gamma \vdash \psi$.

(c) **Chaining.** If $\Gamma \vdash \alpha_i$ for **every** $i = 1, \ldots, m$, and $\alpha_1, \ldots, \alpha_m \vdash \beta$, then $\Gamma \vdash \beta$.

In words: (a) extra premises never hurt; (b) a statement that has been derived may be used as a premise; (c) a sequent already proved can be applied to statements already derived. Part (c) is Definition 2.3.1 for whole derivations, outside every box; inside a box, the expansion of the fold "Why a derived rule adds no power" does the same job.

Chaining and cut are the reason a derived-rule line outside every box is legitimate: the lines it cites lie in no box, so the derivation up to each of them derives it from Γ, and (c) then derives the instance of the sequent's conclusion from Γ; by (b), the rest of the derivation may use that instance.

Idea. Part (a) is read off Definition 1.2.3, which asks only that the premises of the derivation **belong to** the list. For (b) and (c), write the derivations one after another and let the last one cite, instead of its premise α_i, the last line of the derivation of α_i. The only thing to check is the scope rule of Definition 1.4.2: the last line of a derivation lies in no box, so it stays available to everything written after it.

Proof. (a) A derivation witnessing $\Gamma \vdash \psi$ has premises that belong to Γ, hence to Δ, and its last line is ψ. By Definition 1.2.3 it witnesses $\Delta \vdash \psi$. The second sentence is the case where Δ is Γ, φ.

For (b) and (c) we prove one statement that contains both:

$(*)$ If $\Gamma \vdash \alpha_i$ for every $i = 1, \dots, m$, and $\Gamma, \alpha_1, \dots, \alpha_m \vdash \beta$, then $\Gamma \vdash \beta$.

Part (b) is $(*)$ with $m = 1$, $\alpha_1 = \varphi$ and $\beta = \psi$. Part (c) follows from $(*)$, since $\alpha_1, \dots, \alpha_m \vdash \beta$ gives $\Gamma, \alpha_1, \dots, \alpha_m \vdash \beta$ by (a).

To prove $(*)$, let D_i be a derivation of α_i with premises in Γ, for each i, and let E be a derivation of β with premises in $\Gamma, \alpha_1, \dots, \alpha_m$. If the last line of E is one of its premises, then β is a member of Γ, and the one-line derivation whose only line is the premise β shows $\Gamma \vdash \beta$; or β is some α_i, and D_i shows $\Gamma \vdash \beta$. Otherwise, build a list of lines F as follows.

1. First, as premises, every premise of $D_1, \dots, D_m$, and every premise of E that is a member of Γ.
2. Then the lines of D_1 that are not premises, then those of D_2, and so on up to D_m, each with its boxes.
3. Then the lines of E that are not premises, with its boxes.

Renumber the lines, and change each citation as follows. A citation of a premise that is in Γ now cites the premise line of step 1 carrying that statement. A citation in E of a premise α_i that is not in Γ now cites the last line of D_i; that line is not a premise of D_i, because the premises of D_i are in Γ and α_i is not, so it was copied in step 2. Every other citation follows its line to its new number.

We check that F is a derivation (Definition 1.4.2). The premises come first and lie in no box. The boxes are those of the D_i and of E, moved as blocks, so two boxes are still disjoint or nested, each still begins with its assumption, and each line lies in exactly the boxes it lay in before. Hence a line and a box that were available at a citing line are still available there, since the boxes containing each are unchanged. A premise line of step 1 lies in no box and comes first, so it is available at every later line. The last line of D_i lies in no box (it was the last line of a derivation) and comes before every line of E, so it is available at every line of E. Each line carries the same statement and the same justification as before, citing lines with the same statements, so it is an instance of the same rule. Finally, the last line of F is the last line of E, namely β, and it lies in no box. The premises of F are all in Γ. Therefore $\Gamma \vdash \beta$, which proves $(*)$.

■

Quick check. Given Proposition 2.3.1, which part of Lemma 2.3.3 gives $P \to Q,\ \neg Q,\ R \vdash \neg P$ at once? Given also Proposition 2.1.2, which part gives $P \to Q,\ \neg Q \vdash P \to R$?

Solution. Part (a), weakening, with R as the extra premise. Part (c), chaining, with $m = 1$, $\alpha_1 = \neg P$ and $\beta = P \to R$: Proposition 2.3.1 gives $P \to Q,\ \neg Q \vdash \neg P$, and Proposition 2.1.2 with R in place of Q gives $\neg P \vdash P \to R$. (Part (a) and then part (b) do the same.)

2.3.5 Disjunctive syllogism: the impossible case, filed

The second pattern is the impossible case of Section 2.1. Every integer is even or odd; this one is not even; so it is odd. A disjunction, one side refuted, the other side left.

Proposition 2.3.4. Disjunctive Syllogism

$P \vee Q,\ \neg P \vdash Q$.

Idea. This is Example (An impossible case) with a simpler clash. There, the case P produced R by $\to$E, which clashed with $\neg R$. Here the case P clashes with the premise $\neg P$ at once. So: split the premise $P \vee Q$ by $\vee$E, both boxes ending in the current goal Q. The Q box is its own assumption. The P box cannot reach Q honestly, so it reaches a clash and closes by $\bot$E: an impossible case closes itself.

Proof. The following is a derivation with premises $P \vee Q$ and $\neg P$, and last line Q.

1	$P \vee Q$	
2	$\neg P$	
3	$\quad P$	asm, for $\vee$E
4	$\quad \bot$	$\neg$E 3, 2
5	$\quad Q$	$\bot$E 4
6	$\quad Q$	asm, for $\vee$E
7	Q	$\vee$E 1, 3–5, 6–6

The two boxes 3–5 and 6–6 are separate. The second consists of its assumption alone, which is already the statement Q that $\vee$E needs. This proves the proposition. ■

Set the derivation beside that of Example (An impossible case). Both split $P \vee Q$ into the same two boxes, and in both the Q box is one line long. Inside the P box, two lines reached $\bot$ there (lines 5–6) and one line does here (line 4); then $\bot$E closes the box in the same way. From now

on this whole stretch is one line, tagged DS.

The rule as stated refutes the **left** disjunct. To refute the right one, first put the disjunction in the other order; that, too, is now one line, since Example (First derivations with "or") derived $P \vee Q \vdash Q \vee P$. (Exercise B1 derives the mirror image directly.)

Example. Derived rules in use

(a) Show that $P \to Q$, $Q \to R$, $\neg R \vdash \neg P$, and then write the derivation out with primitive rules only.

(b) Show that $P \vee Q$, $\neg P$, $Q \to R \vdash R$.

Solution. Idea. A derived rule is a **premise-directed shortcut**: a line $\varphi \to \psi$ with a line $\neg\psi$ invites MT; a disjunction with one side refuted invites DS. This is the move this section adds to the search for derivations: before opening a box by hand, ask whether a derived rule takes the step at once.

(a) Work backwards from the goal $\neg P$. The premise $P \to Q$ would give it by MT if $\neg Q$ were available. And $\neg Q$ comes by MT from $Q \to R$ and $\neg R$. So MT is used twice, the second time on the conclusion of the first.

1	$P \to Q$
2	$Q \to R$
3	$\neg R$
	⋮
5	$\neg P$

1	$P \to Q$	
2	$Q \to R$	
3	$\neg R$	
	⋮	
4	$\neg Q$	
5	$\neg P$	MT 1, 4

1	$P \to Q$	
2	$Q \to R$	
3	$\neg R$	
4	$\neg Q$	MT 2, 3
5	$\neg P$	MT 1, 4

Line 4 is the instance with φ replaced by Q and ψ by R; line 5 the instance with P and Q. In primitive rules, each MT line becomes its four lines, as in lines 3–6 of the derivation of Proposition 2.3.1: a box on the hypothesis, →E, ¬E, ¬I. This natural expansion has eleven lines.

1	$P \to Q$	
2	$Q \to R$	
3	$\neg R$	
4	$\quad Q$	asm, for $\neg$I
5	$\quad R$	$\to$E 2, 4
6	$\quad \bot$	$\neg$E 5, 3
7	$\neg Q$	$\neg$I 4–6
8	$\quad P$	asm, for $\neg$I
9	$\quad Q$	$\to$E 1, 8
10	$\quad \bot$	$\neg$E 9, 7
11	$\neg P$	$\neg$I 8–10

The literal expansion of the fold "Why a derived rule adds no power" adds a repetition of each cited line and is longer; the natural one cites the available lines directly.

(b) The disjunction $P \vee Q$ has its left side refuted by $\neg P$, so DS gives Q, and then $Q \to R$ fires.

1	$P \vee Q$
2	$\neg P$
3	$Q \to R$
	$\vdots$
5	R

1	$P \vee Q$	
2	$\neg P$	
3	$Q \to R$	
	$\vdots$	
4	Q	
5	R	$\to$E 3, 4

1	$P \vee Q$	
2	$\neg P$	
3	$Q \to R$	
4	Q	DS 1, 2
5	R	$\to$E 3, 4

Expanded, line 4 becomes lines 3–7 of the derivation of Proposition 2.3.4 (five lines), with a case split on line 1 and the case P closed by $\bot$E.

Quick check. Which derived rule of this section, if any, gives a new line in one step from each pair of available lines?

(i) $P \to Q$ and $\neg Q$. (ii) $P \vee Q$ and $\neg P$. (iii) $P \vee Q$ and $Q \to R$. (iv) $P \to \neg Q$ and Q.

Solution.

(i) MT gives $\neg P$. (ii) DS gives Q. (iii) None: nothing is known about the case P. (iv) None in one step: MT needs the negation of the conclusion $\neg Q$, which is $\neg\neg Q$, not Q. A $\neg$I box gives $\neg\neg Q$ from Q first, as in Example (Two negations from one premise), and then MT gives $\neg P$.

2.3.6 Modus tollens on numbers, and the green card

In a mathematical proof, MT is the step "since the conclusion fails, the hypothesis fails". Here it is with numbers.

Example. Modus tollens on numbers

Let x be a real number with $x^2 \leq 4$. Use the implication "if $x > 2$, then $x^2 > 4$" to show that $x \leq 2$, naming the derived rule.

Solution. Take P: "$x > 2$" and Q: "$x^2 > 4$". The implication is $P \to Q$. (It holds because $x > 2 > 0$ gives $x^2 = x \cdot x > 2x > 4$.) The hypothesis $x^2 \leq 4$ is $\neg Q$, since for real numbers $x^2 \leq 4$ and $x^2 > 4$ are exact negations of each other. By MT (Proposition 2.3.1), $\neg P$, that is, $\neg(x > 2)$, which is $x \leq 2$.

The claim of Example (Four cards) (Section 1.3), for one card, is $E \to R$, with E: "this card's number is even" and R: "its color side is red".

The card shows	So we have	What a rule gives	Turn it?
6	E	$\to$E: R; the back must be red	yes: a green back would break the claim
7	$\neg E$	nothing about R: no rule takes $E \to R$ and $\neg E$ to R or to $\neg R$	no
red	R	nothing about E: concluding E would be affirming the consequent	no
green	$\neg R$	MT: $\neg E$; the number must be odd	yes: an even number would break the claim

Turning the green card tests exactly the line MT wrote: an even number there gives E, and with $\neg R$ and $E \to R$, $\to$E and $\neg$E reach $\bot$. Modus tollens is the step people skip.

CP turns the claim into its contrapositive $\neg R \to \neg E$, "a card whose color is not red has an odd number", which carries the claim's content to the green card; Section 2.5 derives the passage

back, so the two say the same. The converse $R \to E$, "a red card has an even number", is a different claim, and the card rule never made it.

Quick check. Which card would MT settle if the claim were "if a card's color side is red, then its number is even"?

Solution. The 7. The claim is $R \to E$, the 7 shows $\neg E$, and MT gives $\neg R$, so its back must not be red: turn it.

↻ **Second reading** · *Modus tollens and testing a claim*

A claim about every integer of a kind makes one prediction per instance. "For every prime p, the number $2^p - 1$ is prime" predicts that 3, 7, 31 and 127 are prime, for $p = 2, 3, 5, 7$, and all four are. Write A for the claim and B for one prediction, so that $A \to B$. The form "from $A \to B$ and B, conclude A" breaks (it is affirming the consequent, Section 1.3), so no number of confirming instances proves the claim. One failed instance refutes it, by MT on that instance: for $p = 11$, $2^{11} - 1 = 2047 = 23 \cdot 89$, so the prediction B fails, and from $A \to B$ and $\neg B$, MT gives $\neg A$. Scientists call a hypothesis that has survived many tests *corroborated*, not proved; a mathematician calls such a claim a conjecture.

2.3.7 How this rule looks in a mathematical proof

Theorem 2.1.3 (Section 2.1) says that for an integer n, if n^2 is even then n is even. Section 2.1 proved it by cases, with the odd case shown impossible and closed by $\bot$E. Here is the same theorem argued as mathematicians usually write it, with the derived rules doing the work. We use the facts (E1)–(E3) of Section 2.1: every integer is even or odd; no integer is both; the square of an odd integer is odd.

Written the way a mathematician would write it, the proof is four sentences: *Suppose n^2 is even. Then n^2 is not odd, since no integer is both. By the contrapositive of (E3), n is not odd. As every integer is even or odd, n is even.* Each sentence after the first is one named step: $\neg$I with (E2); CP with $\to$E; DS, after exchanging the disjuncts of (E1), as line 11 of the derivation below does. The example below writes the same proof with every step tagged, and then as a derivation.

Example. Even square, even number, by contraposition

Let n be an integer. Prove that if n^2 is even, then n is even, using CP and DS, and label each step with its rule.

Solution. Idea. The hypothesis "n^2 is even" hands us nothing to compute with, while "n is odd" hands us $n = 2k + 1$. So we turn (E3) round by CP: a number whose square is not odd is not odd itself. Then (E1), with one side refuted, leaves "even" by DS.

Suppose n^2 is even [assumption for $\to$I]. If n^2 were odd [assumption for $\neg$I], it would be

both even and odd [∧I], contradicting (E2) [¬E]; hence n^2 is not odd [¬I]. By (E3), if n is odd then n^2 is odd, so by contraposition, if n^2 is not odd then n is not odd [CP]. Hence n is not odd [→E]. By (E1), n is even or odd, so n is odd or even [or commutes], by the exchange of disjuncts of Example (First derivations with "or") used as a derived rule. Since n is not odd, n is even [DS]. This proves that if n^2 is even, then n is even [→I].

With E: "n is even", O: "n is odd", S: "n^2 is even", T: "n^2 is odd", and the three facts as premises, the skeleton is this derivation.

1	$E \lor O$	
2	$\neg(S \land T)$	
3	$O \to T$	
4	$\quad S$	asm, for →I
5	$\quad\quad T$	asm, for ¬I
6	$\quad\quad S \land T$	∧I 4, 5
7	$\quad\quad \bot$	¬E 6, 2
8	$\quad \neg T$	¬I 5–7
9	$\quad \neg T \to \neg O$	CP 3
10	$\quad \neg O$	→E 9, 8
11	$\quad O \lor E$	or commutes 1
12	$\quad E$	DS 11, 10
13	$S \to E$	→I 4–12

Line 11 is Example (First derivations with "or") used as a derived rule, so that line 12 refutes the left disjunct, as DS requires.

Second reading · *The two derivations of Even Square, line for line*

Compare this with the twelve-line derivation after Theorem 2.1.3 in Section 2.1. There, line 11 split on (E1) at once, and the odd case, lines 6–10, got T from (E3), built $S \land T$, clashed with (E2), and closed by ⊥E. Here the same work is spread over named steps. The ¬I box on lines 5–7 does the clash with (E2), which there was lines 8–9. Lines 9 and 10, CP and →E, abbreviate a box that assumes O, gets T by (E3), as line 7 did there, and clashes with line 8. Line 12, DS, abbreviates a case split on the disjunction with the case O closed by ⊥E, as lines 5–11 did there. The short version reads the way people talk.

2.3.8 Common errors

Warning. A derived rule may be used only **after** its sequent has been derived, and never inside its own derivation. A "derivation" of Proposition 2.3.2 that justifies one of its lines by

CP proves nothing: it assumes what it sets out to show. Exercise C2 is an example.

Warning. The converse of the contrapositive is a different sequent. CP passes from $P \to Q$ to $\neg Q \to \neg P$; it does not pass from $\neg Q \to \neg P$ back to $P \to Q$. Try it: in a box on P the goal is Q, a letter with no negation, and the natural box, on $\neg Q$, reaches $\bot$ and so closes to $\neg\neg Q$, not Q. Section 2.5 derives the passage back with a new rule.

2.3.9 The rule card

Example. Instances of MT**,** CP **and** DS

A sequent is proved once, for letters, and used for every statement put in their place. So it pays to prove a lemma with the fewest premises and the most general letters it allows: one derivation, written once, serves every row of its rule below.

Rule	Letters replaced	Cited lines	New line
MT	$\varphi := P \lor Q$, $\psi := R \land S$	$(P \lor Q) \to (R \land S)$, $\neg(R \land S)$	$\neg(P \lor Q)$
CP	$\varphi := \neg P$, $\psi := Q \to R$	$\neg P \to (Q \to R)$	$\neg(Q \to R) \to \neg\neg P$
DS	$\varphi := P \to Q$, $\psi := \neg R \lor S$	$(P \to Q) \lor (\neg R \lor S)$, $\neg(P \to Q)$	$\neg R \lor S$

Not an instance. From $(P \to Q) \lor R$ and $\neg P$, DS gives nothing: the refuted line must be the negation of the whole left disjunct, $\neg(P \to Q)$, and $\neg P$ is not that line.

Chapter 1's seven rules are on the card at the end of Section 1.6; here are this chapter's so far.

Rule	Shape	You write	Lean
¬E	$\varphi, \neg\varphi \Longrightarrow \bot$	"…, which contradicts …"	`exact hn hp` (goal `False`)
⊥E	$\bot \Longrightarrow$ any φ	"This case cannot occur." "So there is nothing to prove."	`exfalso`
¬I	box φ … $\bot \Longrightarrow \neg\varphi$	"Suppose P. … This is a contradiction. Hence not P."	`intro hp` (goal `False`)
MT	(derived) $\varphi \to \psi, \neg\psi \Longrightarrow \neg\varphi$	"Since Q fails, so does P."	`mt h hnq`
CP	(derived) $\varphi \to \psi \Longrightarrow \neg\psi \to \neg\varphi$	"We prove the contrapositive: suppose not Q; … so not P."	`mt`

Rule	Shape	You write	Lean
DS	(derived) $\varphi \lor \psi$, $\neg\varphi \Longrightarrow \psi$	"P is impossible, so Q."	`h.resolve_left hnp`

Warning (Not allowed: derived rules). Each step breaks the clause of the reading of Definition 2.3.1 shown.

- "CP on a line of the derivation of Proposition 2.3.2": the sequent is not yet derived where it is used (clause 1; Exercise C2).
- "$\neg P$ by MT from $(P \land Q) \to R$ and $\neg R$": MT writes the negation of the whole hypothesis, $\neg(P \land Q)$ (clause 4).
- "MT citing a line inside a closed box": the scope rule still governs every citation (clause 3).
- "$\neg P$ in one line from $P \to Q$, $Q \to R$, $\neg R$": one instance per line; Example (Derived rules in use) (a) takes two MT lines (clause 4).
- "P by $\neg\neg$E from $\neg\neg P$, here": $\neg\neg$E is derived only in Section 2.5 (clause 1).

Remark (Looking ahead). Chapter 4 makes exact the replacement of letters that a derived rule allows, under the name substitution.

2.3.10 Exercises

2.3.10.1 A. Check your understanding

Exercise. A1

(a) State Definition 2.3.1 in your own words, including what "already derived" and "the same statement" require.

(b) Here is a derivation that uses MT. Rewrite it with primitive rules only.

1	$R \to P$	
2	$\neg P$	
3	$\neg R$	MT 1, 2
4	$\neg R \lor S$	$\lor$I 3

(c) True or false, with a reason: "$\neg Q \to \neg P$ follows from $P \to Q$ by CP, so $P \to Q$ follows from $\neg Q \to \neg P$ by CP."

Solution.

(a) If a sequent has been derived earlier (not inside the derivation that uses it), then in a later derivation we may replace its letters by statements, each letter by one statement everywhere it occurs, and, when available lines hold the resulting instances of its premises, write the resulting instance of its conclusion on one line, citing the sequent and those lines.

(b) Line 3 becomes a box on R, then $\to$E, $\neg$E, and $\neg$I:

1	$R \to P$	
2	$\neg P$	
3	$\quad R$	asm, for $\neg$I
4	$\quad P$	$\to$E 1, 3
5	$\quad \bot$	$\neg$E 4, 2
6	$\neg R$	$\neg$I 3–5
7	$\neg R \lor S$	$\lor$I 6

(c) False. CP is the sequent $P \to Q \vdash \neg Q \to \neg P$. Applied to $\neg Q \to \neg P$, with P replaced by $\neg Q$ and Q by $\neg P$, it gives $\neg\neg P \to \neg\neg Q$, not $P \to Q$. The passage from $\neg Q \to \neg P$ to $P \to Q$ is the converse, a different sequent; Section 2.5 derives it with a new rule.

Exercise. A2: Drill

Each part takes a line or two. The parts run from single steps, through a missing line and a justification column, to whole derivations. Let x be a real number.

(a) Lines 1 $(A \land B) \to C$, 2 $\neg C$. What does "MT 1, 2" write?
(b) Lines 1 $\neg P \lor Q$, 2 $\neg\neg P$. What does "DS 1, 2" write?
(c) What does CP write from $(x > 0) \to (x^2 > 0)$?
(d) Line 3 is $\neg(P \lor Q)$ and line 5 is $\neg R \to (P \lor Q)$. What does "MT 5, 3" write?
(e) Written out with primitive rules, citing available lines directly, how many lines does one MT line become?
(f) Line 8 is $\neg(R \land S)$, justified "MT 2, 6", and line 6 is $\neg T$. What is line 2?
(g) Line 9 is $R \to S$, justified "DS 4, 7", and line 7 is $\neg(P \land Q)$. What is line 4?
(h) Lines 1–4 are the premises $P \lor (Q \to R)$, $\neg P$, $S \to Q$, $\neg R$. Justify 5 $Q \to R$; 6 $\neg Q$; 7 $\neg S$.
(i) Lines 1 $P \to \neg Q$, 2 Q. Does MT apply?
(j) Lines 1 $P \lor Q$, 2 $\neg Q$. Does DS, as stated, apply?
(k) May a derivation of Proposition 2.3.2 use CP on one of its lines?
(l) Find the illegal line: 1 $P \to Q$; 2 $\neg P$; 3 $\neg Q$, MT 1, 2.
(m) Derive $P \to (Q \land R),\ \neg R \vdash \neg P$.
(n) Derive $P \lor Q,\ \neg P,\ \neg Q \vdash R$.
(o) Derive $\neg P \to Q,\ \neg Q \vdash \neg\neg P$ in one line.

Solution.

(a) $\neg(A \land B)$.
(b) Q (the letter P of DS replaced by $\neg P$).
(c) $\neg(x^2 > 0) \to \neg(x > 0)$, that is, $(x^2 \leq 0) \to (x \leq 0)$.
(d) $\neg\neg R$, with $\varphi := \neg R$ and $\psi := P \lor Q$.

(e) Four: an assumption, $\to$E, $\neg$E, $\neg$I (the figure).

(f) $(R \wedge S) \to T$.

(g) $(P \wedge Q) \vee (R \to S)$.

(h) 5 DS 1, 2; 6 MT 5, 4; 7 MT 3, 6.

(i) Not yet: MT needs $\neg\neg Q$. Write it from Q by a $\neg$I box first; then MT gives $\neg P$.

(j) No: DS refutes the left disjunct. Use the other form (Exercise B1 (a)) or exchange the disjuncts first.

(k) No: a derived rule may be used only after its sequent has been derived.

(l) Line 3: MT needs the negation of the conclusion, $\neg Q$, and gives $\neg P$; line 3 runs it backwards, from $\neg P$ to $\neg Q$. The form breaks at P false, Q true.

(m) 3 P, assumption; 4 $Q \wedge R$, $\to$E 1, 3; 5 R, $\wedge$E 4; 6 $\bot$, $\neg$E 5, 2; 7 $\neg P$, $\neg$I 3–6. (MT does not match: $\neg(Q \wedge R)$ is not available.)

(n) 4 Q, DS 1, 2; 5 $\bot$, $\neg$E 4, 3; 6 R, $\bot$E 5.

(o) 3 $\neg\neg P$, MT 1, 2.

2.3.10.2 B. Practice

Exercise. B1: The other disjunctive syllogism, used later

(a) Show that $P \vee Q,\ \neg Q \vdash P$ by a derivation from the primitive rules.

(b) Let x be a real number with $x^2 = 1$ and $x > 0$. You may use that $x^2 = 1$ implies $x = 1$ or $x = -1$. Deduce that $x = 1$, naming the result you use.

Solution.

(a) As in Proposition 2.3.4, with the roles of the boxes exchanged: now the case Q is the impossible one.

1	$P \vee Q$	
2	$\neg Q$	
3	$\quad P$	asm, for $\vee$E
4	$\quad Q$	asm, for $\vee$E
5	$\quad \bot$	$\neg$E 4, 2
6	$\quad P$	$\bot$E 5
7	P	$\vee$E 1, 3–3, 4–6

(b) Since $x^2 = 1$, we have $x = 1$ or $x = -1$. Since $x > 0$ and $-1 < 0$, we have $x \neq -1$. By part (a), with P: "$x = 1$" and Q: "$x = -1$", used as a derived rule (Definition 2.3.1), we conclude $x = 1$.

Exercise. B2: Putting a derivation in order

Here are nine lines, each a statement with its justification. Seven of them, in the right order and with the right boxes, form a derivation of $P \to Q \vdash \neg Q \to \neg P$ from the primitive rules.

Give the order, and say where each box opens and closes. Then say why each of the other two lines cannot be used.

(i) $\neg Q \to \neg P$, by $\to$I on the box from $\neg Q$ to $\neg P$.
(ii) Q, by $\to$E from $P \to Q$ and P.
(iii) $P \to Q$, premise.
(iv) $\neg P$, by $\neg$I on the box from P to $\bot$.
(v) $\neg Q$, assumption.
(vi) $\bot$, by $\neg$E from Q and $\neg Q$.
(vii) P, assumption.
(viii) $\neg P$, assumption, opening the box for the $\neg$I that proves $\neg P$.
(ix) Q, by $\to$E from $P \to Q$ and the assumption P of (vii), written directly below (iv).

Solution. The order is (iii), (v), (vii), (ii), (vi), (iv), (i), with one box opened by (v) and closed after (iv), and a second box inside it, opened by (vii) and closed after (vi). This is the derivation of Proposition 2.3.2. It is the only admissible order. The premise comes first. (i) needs a box from $\neg Q$ to $\neg P$, so (v) opens a box whose last line is (iv). (iv) needs a box from P to $\bot$ inside it, so (vii) opens a box whose last line is (vi). Inside that box, (ii) needs (vii), and (vi) needs (ii) and (v).

(viii) cannot be used. $\neg$I proves $\neg P$ from a box opened by P, the statement being negated (Definition 2.2.1). A box opened by $\neg P$ and ending in $\bot$ would give $\neg\neg P$. (ix) cannot be used where it is placed. Directly below (iv), the box opened by (vii) has closed, so the assumption P is not available (Definition 1.4.2). Placed inside that box, the same line is (ii).

Exercise. B3: Two refutations

A search for $(P \vee Q) \to R,\ \neg R \vdash \neg P \wedge \neg Q$ has reached the stage below. You may use the derived rules of this section. *Name the move, then complete it.*

1	$(P \vee Q) \to R$	
2	$\neg R$	
3	$\neg(P \vee Q)$	MT 1, 2
	$\vdots$	
12	$\neg P \wedge \neg Q$	

Solution. The premises are an implication and the negation of its conclusion, which invites MT: it gives $\neg(P \vee Q)$. The goal is a conjunction, so we need $\neg P$ and $\neg Q$, each a negation, each by a $\neg$I box. In the box of P, the clash is manufactured: $\vee$I gives $P \vee Q$, which clashes with line 3. The box of Q is the same.

1	$(P \vee Q) \to R$	
2	$\neg R$	
3	$\neg(P \vee Q)$	MT 1, 2
4	$\quad P$	asm, for $\neg$I
5	$\quad P \vee Q$	$\vee$I 4
6	$\quad \bot$	$\neg$E 5, 3
7	$\neg P$	$\neg$I 4–6
8	$\quad Q$	asm, for $\neg$I
9	$\quad P \vee Q$	$\vee$I 8
10	$\quad \bot$	$\neg$E 9, 3
11	$\neg Q$	$\neg$I 8–10
12	$\neg P \wedge \neg Q$	$\wedge$I 7, 11

Line 3 is MT with φ replaced by $P \vee Q$ and ψ by R.

Exercise. B4: Dilemmas as derived rules, used later

(a) Section 1.6 derived the constructive dilemma, Example (Constructive dilemma). Use it as a derived rule, tagged CD, together with Example (First derivations with "or"), to derive $(P \to R) \wedge (Q \to R),\ P \vee Q \vdash R$ in four lines.

(b) Derive the destructive dilemma $(P \to R) \wedge (Q \to S),\ \neg R \vee \neg S \vdash \neg P \vee \neg Q$, using MT.

(c) Derive $\neg(P \wedge Q) \vdash P \to \neg Q$: "$P$ rules out Q", the form in which the traditional *modus ponendo tollens* is used. Then break "from $\neg(P \wedge Q)$ and $\neg P$, conclude Q".

(d) Let n be an integer. Use (b) to show: if n is odd or n is not a multiple of 3, then n is not a multiple of 4 or not a multiple of 9. Take as premises that a multiple of 4 is even and a multiple of 9 is a multiple of 3, and read "odd" as "not even".

Solution.

(a) Line 3 is Example (Constructive dilemma), $(P \to R) \wedge (Q \to S),\ P \vee Q \vdash R \vee S$, with S replaced by R. Line 4 is the second sequent of Example (First derivations with "or"), $P \vee P \vdash P$, with P replaced by R.

1	$(P \to R) \wedge (Q \to R)$	
2	$P \vee Q$	
3	$R \vee R$	CD 1, 2
4	R	or merges 3

(b) Split the premise $\neg R \vee \neg S$; in each case MT refutes one hypothesis, and $\vee$I builds the goal.

1	$(P \to R) \land (Q \to S)$	
2	$\neg R \lor \neg S$	
3	$P \to R$	$\land$E 1
4	$Q \to S$	$\land$E 1
5	$\quad \neg R$	asm, for $\lor$E
6	$\quad \neg P$	MT 3, 5
7	$\quad \neg P \lor \neg Q$	$\lor$I 6
8	$\quad \neg S$	asm, for $\lor$E
9	$\quad \neg Q$	MT 4, 8
10	$\quad \neg P \lor \neg Q$	$\lor$I 9
11	$\neg P \lor \neg Q$	$\lor$E 2, 5–7, 8–10

(c) The goal is an implication, so open a box on P; its goal $\neg Q$ is a negation, so open a $\neg$I box on Q inside it.

1	$\neg(P \land Q)$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad Q$	asm, for $\neg$I
4	$\quad\quad P \land Q$	$\land$I 2, 3
5	$\quad\quad \bot$	$\neg$E 4, 1
6	$\quad \neg Q$	$\neg$I 3–5
7	$P \to \neg Q$	$\to$I 2–6

The look-alike breaks with P: "$x > 5$" and Q: "$x > 3$", at $x = 0$: $\neg(P \land Q)$ and $\neg P$ are true and Q is false, so the counterexample set $\neg(P \land Q)$, $\neg P$, $\neg Q$ is true. *Modus ponendo tollens* uses a negated conjunction and **affirms** one conjunct; DS uses a disjunction and **refutes** one disjunct; mixing the two is the error of the look-alike.

(d) Write F: "n is a multiple of 4", E: "n is even", N: "n is a multiple of 9", T: "n is a multiple of 3". The premises are $(F \to E) \land (N \to T)$ (since $4k = 2(2k)$ and $9k = 3(3k)$), and the hypothesis is $\neg E \lor \neg T$. Open a box on the hypothesis; inside it, (b) as a derived rule, with P, Q, R, S replaced by F, N, E, T, gives $\neg F \lor \neg N$; then $\to$I closes the box on $(\neg E \lor \neg T) \to (\neg F \lor \neg N)$. For instance $n = 7$ is odd, and not a multiple of 4.

Marking scheme. (10 points)

- 2: (a), CD with the replacement $S \coloneqq R$ named, then $P \lor P \vdash P$.
- 3: (b), MT in each case box, both boxes ending in $\neg P \lor \neg Q$.
- 3: (c), the nested $\neg$I box inside the $\to$I box (2), and a correct break of the look-alike (1).
- 2: (d), the letters, the premises, and the use of (b) inside a $\to$I box.

Fatal error (0 for (b)): a case box citing a line of the other case box.

Exercise. B5: Derive or break

For each claim, derive it with the rules so far, including MT, CP, DS and the sequents derived in this book so far, or break it by statements about a real number x that make its counterexample set true at one value of x.

(a) $P \lor Q,\ Q \vdash \neg P$.
(b) $(P \lor R) \to \neg Q,\ Q \vdash \neg R$.
(c) $\neg P \to Q \vdash \neg Q \to \neg P$.
(d) $P \lor (Q \land R),\ \neg P \vdash R$.
(e) $(P \land Q) \to R,\ \neg R,\ P \vdash \neg Q$.
(f) $P \to (Q \lor R),\ \neg Q \vdash P \to R$.
(g) $P \to (Q \lor R),\ \neg Q \vdash R$.
(h) $(P \to Q) \land (R \to S),\ \neg Q \lor \neg S \vdash \neg P \land \neg R$.

Solution.

(a) Break. Take P: "$x > 0$" and Q: "$x > 1$". At $x = 2$, $P \lor Q$ and Q are true and P is true, so the counterexample set (Definition 1.1.3) $P \lor Q$, Q, $\neg\neg P$ is true. This is affirming a disjunct (Section 1.6).

(b) Derive. The goal is a negation: open a box on R, and manufacture the premise's hypothesis by $\lor$I.

1	$(P \lor R) \to \neg Q$	
2	Q	
3	$\quad R$	asm, for $\neg$I
4	$\quad P \lor R$	$\lor$I 3
5	$\quad \neg Q$	$\to$E 1, 4
6	$\quad \bot$	$\neg$E 2, 5
7	$\neg R$	$\neg$I 3–6

(c) Break. Take P: "$x > 0$" and Q: "$x > 5$". At $x = 1$, $\neg P \to Q$ is true (its hypothesis fails), $\neg Q$ is true and $\neg P$ is false, so the counterexample set $\neg P \to Q$, $\neg(\neg Q \to \neg P)$ is true. (CP on $\neg P \to Q$ gives $\neg Q \to \neg\neg P$, not this.)

(d) Derive.

1	$P \lor (Q \land R)$	
2	$\neg P$	
3	$Q \land R$	DS 1, 2
4	R	$\land$E 3

(e) Derive. MT refutes the conjunction; then a $\neg$I box on Q rebuilds it from the premise P.

1	$(P \wedge Q) \to R$	
2	$\neg R$	
3	P	
4	$\neg(P \wedge Q)$	MT 1, 2
5	$\quad Q$	asm, for $\neg$I
6	$\quad P \wedge Q$	$\wedge$I 3, 5
7	$\quad \bot$	$\neg$E 6, 4
8	$\neg Q$	$\neg$I 5–7

(f) Derive.

1	$P \to (Q \vee R)$	
2	$\neg Q$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q \vee R$	$\to$E 1, 3
5	$\quad R$	DS 4, 2
6	$P \to R$	$\to$I 3–5

(g) Break. Take P: "$x > 5$", Q: "$x > 3$", R: "$x > 4$". The premise $P \to (Q \vee R)$ is true for every x. At $x = 0$, $\neg Q$ is true and R is false, so the counterexample set $P \to (Q \vee R)$, $\neg Q$, $\neg R$ is true.

(h) Break. Take P: "$x > 5$", Q: "$x > 9$", R: "$x > 0$", S: "$x > -1$". At $x = 1$, $P \to Q$ is true (its hypothesis fails) and $R \to S$ is true (for every x), $\neg Q \vee \neg S$ is true since $\neg Q$ is, and $\neg P \wedge \neg R$ is false, since R is true. So the counterexample set $(P \to Q) \wedge (R \to S)$, $\neg Q \vee \neg S$, $\neg(\neg P \wedge \neg R)$ is true.

(i) and (c) are the look-alikes of DS and of CP; (g) drops the hypothesis that (f) keeps in its box; (h) is the destructive dilemma of Exercise B4 (b) with the $\vee$ of its conclusion made $\wedge$.

Marking scheme. (16 points) For each claim: 1 for the verdict; 1 for a derivation every line of which checks, or for a choice of statements and one value of x at which **every** member of the counterexample set is true. Fatal errors (0 for the claim): a "break" at which a premise fails; RAA or $\neg\neg$E in a derivation.

2.3.10.3 C. Going deeper

Exercise. C1: Two premises that cannot both fail, used later

(a) Show that $P \to Q$, $\neg P \to R$, $\neg Q$, $\neg R \vdash \bot$, using MT twice.

(b) Hence show that $P \to Q$, $\neg P \to R \vdash \neg(\neg Q \wedge \neg R)$ with the rules so far.

(c) In words, the conclusion of (b) says that Q and R do not both fail. Explain why a derivation with the rules so far naturally ends at $\neg(\neg Q \wedge \neg R)$ rather than at $Q \vee R$.

Solution.

(a) The first MT has $\varphi = P$, $\psi = Q$. The second has $\varphi = \neg P$, $\psi = R$, so it gives the negation of $\neg P$, which is $\neg\neg P$. Lines 5 and 6 then clash.

1	$P \to Q$	
2	$\neg P \to R$	
3	$\neg Q$	
4	$\neg R$	
5	$\neg P$	MT 1, 3
6	$\neg\neg P$	MT 2, 4
7	$\bot$	¬E 5, 6

Line 7 cites $\neg P$ and then $\neg\neg P$: the unnegated line first, with φ of ¬E replaced by $\neg P$.

(b) The goal is a negation, so open a box on $\neg Q \wedge \neg R$ and aim at $\bot$. Inside, ∧E supplies the two negated premises of (a), and the rest is (a).

1	$P \to Q$	
2	$\neg P \to R$	
3	$\quad \neg Q \wedge \neg R$	asm, for ¬I
4	$\quad \neg Q$	∧E 3
5	$\quad \neg R$	∧E 3
6	$\quad \neg P$	MT 1, 4
7	$\quad \neg\neg P$	MT 2, 5
8	$\quad \bot$	¬E 6, 7
9	$\neg(\neg Q \wedge \neg R)$	¬I 3–8

(c) To reach $Q \vee R$ by ∨I we would need Q or R on its own. The premises give Q once P is known and R once $\neg P$ is known, and nothing available says which of P, $\neg P$ holds. What the rules so far can do is refute the opposite, $\neg Q \wedge \neg R$, by a ¬I box, and that gives the negation in (b). The step from (b) to $Q \vee R$ uses the rule of Section 2.5.

Exercise. C2: A circular derivation

A student offers this derivation of $P \to Q \vdash \neg Q \to \neg P$.

1	$P \to Q$	
2	$\quad \neg Q$	asm, for →I
3	$\quad \neg Q \to \neg P$	CP 1
4	$\quad \neg P$	→E 3, 2
5	$\neg Q \to \neg P$	→I 2–4

(a) Every line is an instance of the rule it names. What is wrong?

(b) Give a correct derivation. May lines 3 and 4 be replaced by the one line "$\neg P$, MT 1, 2"?

Solution.

(a) Line 3 uses CP, the derived rule whose sequent, $P \to Q \vdash \neg Q \to \neg P$, is the very one being derived. Definition 2.3.1 allows a sequent to be used only when it has **already been derived**; here it has not, so line 3 has no justification. (The whole derivation is also pointless: line 3 alone is already the goal.)

(b) The derivation of Proposition 2.3.2, from the primitive rules:

1	$P \to Q$	
2	$\quad \neg Q$	asm, for $\to$I
3	$\quad\quad P$	asm, for $\neg$I
4	$\quad\quad Q$	$\to$E 1, 3
5	$\quad\quad \bot$	$\neg$E 4, 2
6	$\quad \neg P$	$\neg$I 3–5
7	$\neg Q \to \neg P$	$\to$I 2–6

Yes: lines 3–4 become the one line "$\neg P$, MT 1, 2" inside the box of $\neg Q$, and line 5 becomes $\neg Q \to \neg P$, $\to$I 2–3. This is legitimate, because Proposition 2.3.1 was derived from the primitive rules, without CP. What matters is that the chain of derived rules ends in primitive rules, without a loop.

Exercise. C3: Exit check: labeling a proof

Here is a proof that for an integer n, if n^2 is odd, then n is odd.

> Suppose n^2 is odd. If n were even, $n = 2k$, then $n^2 = 2(2k^2)$ would be even; so, by contraposition, since n^2 is not even (it is odd, and no integer is both), n is not even. As every integer is even or odd, n is odd.

Label each sentence, or part of a sentence, with the rule or rules it uses, primitive or derived. Take as premises the facts (E1) and (E2) of Section 2.1 and the school computation "if n is even, then n^2 is even".

Solution. Write E: "n is even", O: "n is odd", S: "n^2 is even", T: "n^2 is odd".

- "Suppose n^2 is odd": the assumption T of the box for $\to$I, whose goal is O.
- "If n were even, $n = 2k$, then $n^2 = 2(2k^2)$ would be even": the implication $E \to S$, a premise (the school computation). Reading the sentence as a $\to$I box that proves $E \to S$, assuming E and computing, is equally correct.
- "n^2 is not even (it is odd, and no integer is both)": $\neg S$, by $\neg$I: a box on S builds $S \wedge T$ by $\wedge$I and clashes with (E2), $\neg(S \wedge T)$, by $\neg$E.
- "by contraposition ... n is not even": CP on $E \to S$ gives $\neg S \to \neg E$, and $\to$E with

$\neg S$ gives $\neg E$. (One MT line does both.)
- "As every integer is even or odd, n is odd": DS on (E1), $E \vee O$, with the left side E refuted, gives O.
- The frame as a whole: $\to$I, closing the box of T on O, gives $T \to O$.

As a derivation:

1	$E \vee O$	
2	$E \to S$	
3	$\neg(S \wedge T)$	
4	$\quad T$	asm, for $\to$I
5	$\quad\quad S$	asm, for $\neg$I
6	$\quad\quad S \wedge T$	$\wedge$I 5, 4
7	$\quad\quad \bot$	$\neg$E 6, 3
8	$\quad \neg S$	$\neg$I 5–7
9	$\quad \neg S \to \neg E$	CP 2
10	$\quad \neg E$	$\to$E 9, 8
11	$\quad O$	DS 1, 10
12	$T \to O$	$\to$I 4–11

The computation is right: $(2k)^2 = 4k^2 = 2(2k^2)$, and $2k^2$ is an integer.

Marking scheme. (10 points)

- 2: identifies the frame as $\to$I, with the box opened on the hypothesis "n^2 is odd".
- 2: identifies "n^2 is not even" as $\neg$I, naming the clash (with "no integer is both", by $\neg$E).
- 3: identifies CP (then $\to$E), or MT, applied to "if n is even, then n^2 is even", with that antecedent and consequent. Reading the "if n were even" sentence as a premise or as a $\to$I box that proves the implication earns these points equally.
- 2: identifies DS on "every integer is even or odd", naming "even" as the refuted side.
- 1: names the closing $\to$I.

Fatal errors (0 for the step): naming anything but $\neg$I for the "not even" step (it concludes a negation from a box on the unnegated statement); applying CP to the converse, "if n^2 is even, then n is even".

2.4 Iff, Equivalence, and What a Definition Says

One connective is still without rules: "if and only if". Theorems say "n is odd if and only if n^2 is odd", and every definition, though written with a plain "if", is used as an "if and only if". Section 1.1 warned that "if" means different things in a theorem and in a definition; this section explains why. The one new thing here is the pair of rules for $\leftrightarrow$; the rest of the section

(provable equivalence, the layout of an "iff" proof, the convention about definitions, chains of equivalences) is what the rules are used for.

2.4.1 The rules for "if and only if"

"P if and only if Q" combines "P if Q", which is $Q \to P$, with "P only if Q", which is $P \to Q$, as Section 1.1 explained. A proof of it travels both ways; a use of it, either way. An implication is proved by a box, opened on the hypothesis and closed on the conclusion (Definition 1.4.1); an "if and only if" needs two such boxes.

An "if and only if" is two implications, and proving it takes two proofs.

Definition 2.4.1. Rules for the Biconditional

i	φ	asm, for $\leftrightarrow$I
	$\vdots$	
j	ψ	
k	ψ	asm, for $\leftrightarrow$I
	$\vdots$	
l	φ	
m	$\varphi \leftrightarrow \psi$	$\leftrightarrow$I i–j, k–l

i	$\varphi \leftrightarrow \psi$	
j	φ	
k	ψ	$\leftrightarrow$E i, j

i	$\varphi \leftrightarrow \psi$	
j	ψ	
k	φ	$\leftrightarrow$E i, j

The first schema is the rule $\leftrightarrow$I (**iff-introduction**); its two assumptions are discharged when their boxes close, and each box's last line (ψ in the first, φ in the second) is a line **of that box itself**, not of a box inside it. The two schemas below it are the two forms of $\leftrightarrow$E (**iff-elimination**).

As in every schema (Definition 1.2.1), φ and ψ stand for any statements, replaced consistently, and i, j, k, l, m for line numbers. Read the rules clause by clause.

1. **Two boxes.** $\leftrightarrow$I needs two available boxes: one opened by φ and ending in ψ, the other opened by ψ and ending in φ, each last line a line of its own box. Each is exactly the box $\to$I would need for one direction. Their order does not matter, but both must be there: a single box from φ to ψ proves $\varphi \to \psi$ and nothing more. Each box's assumption is printed "asm, for $\leftrightarrow$I"; in a written proof the two boxes are the paragraphs ($\Rightarrow$) and ($\Leftarrow$).
2. **Separate boxes.** The boxes are siblings, not nested, so nothing in the second may cite a line of the first (the scope rule of Definition 1.4.2). Each direction is a proof on its own.
3. **Citation.** The justification cites both boxes as ranges, as in $\leftrightarrow$I 2–4, 5–7.
4. **Elimination, either way.** From $\varphi \leftrightarrow \psi$ and φ on available lines, $\leftrightarrow$E writes ψ; from $\varphi \leftrightarrow \psi$ and ψ, it writes φ. It is modus ponens in whichever direction we need, citing the biconditional first.

In the Givens/Goal table, $\leftrightarrow$I splits a row in two: φ added to the Givens with the goal ψ, and

ψ added with the goal φ.

2.4.2 First derivations with iff

The section's goal-directed move: **a goal $\varphi \leftrightarrow \psi$ opens two boxes**, one on each side, and each box's goal is the other side.

Example. An iff read both ways

(a) Show that $P \leftrightarrow Q \vdash Q \leftrightarrow P$.

(b) Show that $P \leftrightarrow Q,\ \neg Q \vdash \neg P$.

Solution.

(a) *Idea.* The goal is an iff, so the last line is $\leftrightarrow$I, and we need a box from Q to P and a box from P to Q. Each box's goal is one $\leftrightarrow$E away: the premise crosses from its assumption to the other side.
The goal at the foot (left), then its two boxes (right).

1	$P \leftrightarrow Q$	
⋮		
6	$Q \leftrightarrow P$	

1	$P \leftrightarrow Q$	
2	$\quad Q$	asm, for $\leftrightarrow$I
⋮		
3	$\quad P$	
4	$\quad P$	asm, for $\leftrightarrow$I
⋮		
5	$\quad Q$	
6	$Q \leftrightarrow P$	$\leftrightarrow$I 2–3, 4–5

The finished derivation:

1	$P \leftrightarrow Q$	
2	$\quad Q$	asm, for $\leftrightarrow$I
3	$\quad P$	$\leftrightarrow$E 1, 2
4	$\quad P$	asm, for $\leftrightarrow$I
5	$\quad Q$	$\leftrightarrow$E 1, 4
6	$Q \leftrightarrow P$	$\leftrightarrow$I 2–3, 4–5

Line 3 uses the second form of $\leftrightarrow$E (right to left), line 5 the first.

(b) *Idea.* The goal $\neg P$ is a negation, so open a box on P and aim at $\bot$ (Section 2.2). Inside, the premise $P \leftrightarrow Q$ turns P into Q, and Q clashes with $\neg Q$.

1	$P \leftrightarrow Q$	
2	$\neg Q$	
3	$\quad P$	asm, for $\neg$I
4	$\quad Q$	$\leftrightarrow$E 1, 3
5	$\quad \bot$	$\neg$E 4, 2
6	$\neg P$	$\neg$I 3–5

This is the derivation of Proposition 2.3.1 with $\leftrightarrow$E on line 4: if two statements are equivalent and one fails, so does the other.

The rules say nothing that the two implications do not say. The next result makes this exact.

Proposition 2.4.1. An Iff Is Two Implications

$P \leftrightarrow Q \vdash (P \to Q) \wedge (Q \to P)$, and $(P \to Q) \wedge (Q \to P) \vdash P \leftrightarrow Q$.

Idea. Each goal chooses its rule. For the first, the goal is a conjunction of two implications: $\wedge$I needs both, and each is built in its own $\to$I box, crossing the premise with $\leftrightarrow$E inside. For the second, the goal is an iff: open its two boxes. The premise is a conjunction, so take it apart first (premise-directed), and cross each box with one of its halves by $\to$E.

First claim: the two implications as goals (left), then their boxes (right).

1	$P \leftrightarrow Q$	
	$\vdots$	
4	$P \to Q$	
	$\vdots$	
7	$Q \to P$	
8	$(P \to Q) \wedge (Q \to P)$	$\wedge$I 4, 7

1	$P \leftrightarrow Q$	
2	$\quad P$	asm, for $\rightarrow$I
	$\quad \vdots$	
3	$\quad Q$	
4	$P \rightarrow Q$	$\rightarrow$I 2–3
5	$\quad Q$	asm, for $\rightarrow$I
	$\quad \vdots$	
6	$\quad P$	
7	$Q \rightarrow P$	$\rightarrow$I 5–6
8	$(P \rightarrow Q) \wedge (Q \rightarrow P)$	$\wedge$I 4, 7

Second claim: the premise taken apart (left), then the two boxes (right).

1	$(P \rightarrow Q) \wedge (Q \rightarrow P)$	
2	$P \rightarrow Q$	$\wedge$E 1
3	$Q \rightarrow P$	$\wedge$E 1
	$\vdots$	
8	$P \leftrightarrow Q$	

1	$(P \rightarrow Q) \wedge (Q \rightarrow P)$	
2	$P \rightarrow Q$	$\wedge$E 1
3	$Q \rightarrow P$	$\wedge$E 1
4	$\quad P$	asm, for $\leftrightarrow$I
	$\quad \vdots$	
5	$\quad Q$	
6	$\quad Q$	asm, for $\leftrightarrow$I
	$\quad \vdots$	
7	$\quad P$	
8	$P \leftrightarrow Q$	$\leftrightarrow$I 4–5, 6–7

Proof. For the first claim, the following is a derivation with premise $P \leftrightarrow Q$ and last line $(P \rightarrow Q) \wedge (Q \rightarrow P)$.

1	$P \leftrightarrow Q$	
2	$\quad P$	asm, for $\rightarrow$I
3	$\quad Q$	$\leftrightarrow$E 1, 2
4	$P \rightarrow Q$	$\rightarrow$I 2–3
5	$\quad Q$	asm, for $\rightarrow$I
6	$\quad P$	$\leftrightarrow$E 1, 5
7	$Q \rightarrow P$	$\rightarrow$I 5–6
8	$(P \rightarrow Q) \wedge (Q \rightarrow P)$	$\wedge$I 4, 7

For the second claim, the following is a derivation with premise $(P \rightarrow Q) \wedge (Q \rightarrow P)$ and last line $P \leftrightarrow Q$.

1	$(P \rightarrow Q) \wedge (Q \rightarrow P)$	
2	$P \rightarrow Q$	$\wedge$E 1
3	$Q \rightarrow P$	$\wedge$E 1
4	$\quad P$	asm, for $\leftrightarrow$I
5	$\quad Q$	$\rightarrow$E 2, 4
6	$\quad Q$	asm, for $\leftrightarrow$I
7	$\quad P$	$\rightarrow$E 3, 6
8	$P \leftrightarrow Q$	$\leftrightarrow$I 4–5, 6–7

Lines 2 and 3 lie in no box, so both boxes may cite them. This proves both claims.

■

↻ **Second reading** · *Why iff keeps its own rules*

So $\leftrightarrow$ could have been an abbreviation for $(P \rightarrow Q) \wedge (Q \rightarrow P)$. We keep it as a connective with its own rules, so that every connective is proved by an introduction rule and used by an elimination rule.

Example. Negating both sides of an iff

Show that $P \leftrightarrow Q \vdash \neg P \leftrightarrow \neg Q$.

Solution. Idea. The goal is an iff, so there are two boxes, one opened by $\neg P$ and reaching $\neg Q$, the other the reverse. Each box's goal is a negation, so each contains a $\neg$I box: the goal-directed move applied twice, one inside the other. Inside the inner box, $\leftrightarrow$E crosses from one side of the premise to the other, and the result clashes with the outer assumption. The second half is part (b) of Example (An iff read both ways), run inside a box; the first is

its mirror image, with the other form of ↔E.

The outer boxes (left), then a ¬I box inside each (right).

1	$P \leftrightarrow Q$	
2	$\quad \neg P$	asm, for ↔I
	$\quad \vdots$	
6	$\quad \neg Q$	
7	$\quad \neg Q$	asm, for ↔I
	$\quad \vdots$	
11	$\quad \neg P$	
12	$\neg P \leftrightarrow \neg Q$	↔I 2–6, 7–11

1	$P \leftrightarrow Q$	
2	$\quad \neg P$	asm, for ↔I
3	$\quad\quad Q$	asm, for ¬I
	$\quad\quad \vdots$	
5	$\quad\quad \bot$	
6	$\quad \neg Q$	¬I 3–5
7	$\quad \neg Q$	asm, for ↔I
8	$\quad\quad P$	asm, for ¬I
	$\quad\quad \vdots$	
10	$\quad\quad \bot$	
11	$\quad \neg P$	¬I 8–10
12	$\neg P \leftrightarrow \neg Q$	↔I 2–6, 7–11

The finished derivation:

1	$P \leftrightarrow Q$	
2	$\quad \neg P$	asm, for ↔I
3	$\quad\quad Q$	asm, for ¬I
4	$\quad\quad P$	↔E 1, 3
5	$\quad\quad \bot$	¬E 4, 2
6	$\quad \neg Q$	¬I 3–5
7	$\quad \neg Q$	asm, for ↔I
8	$\quad\quad P$	asm, for ¬I
9	$\quad\quad Q$	↔E 1, 8
10	$\quad\quad \bot$	¬E 9, 7
11	$\quad \neg P$	¬I 8–10
12	$\neg P \leftrightarrow \neg Q$	↔I 2–6, 7–11

The two halves are mirror images.

Quick check. In the derivation of Example (Negating both sides of an iff), which line does line 5 cite as $\neg\varphi$, and why may it?

Solution. Line 2, $\neg P$, the assumption of the outer box 2–6, which is still open at line 5.

The derivation uses only the rules of this section and of Sections 2.1–2.2. Its converse, $\neg P \leftrightarrow \neg Q \vdash P \leftrightarrow Q$, is also derivable, with the rule of Section 2.5: in the box of P it must produce the **positive** statement Q, and the natural box, on $\neg Q$, closes to $\neg\neg Q$, not Q.

Example (Negating both sides of an iff) negated both sides of an iff. Can one side be the negation of the other, a statement equivalent to its own negation?

Lemma 2.4.2. No Statement Is Equivalent to Its Negation

$\vdash \neg(P \leftrightarrow \neg P)$.

Idea. The goal is a negation: open a $\neg$I box on $P \leftrightarrow \neg P$, with goal $\bot$. $\leftrightarrow$E needs one side of the iff on a line, and neither is. So make $\neg P$ a subgoal, by a $\neg$I box on P: inside it $\leftrightarrow$E gives $\neg P$, which clashes with P. With $\neg P$ in hand, $\leftrightarrow$E the other way gives P, and the two clash again.

The outer box (left), then the subgoal $\neg P$ with its box (right).

1	$P \leftrightarrow \neg P$	asm, for $\neg$I
⋮	⋮	
7	$\bot$	
8	$\neg(P \leftrightarrow \neg P)$	$\neg$I 1–7

1	$P \leftrightarrow \neg P$	asm, for $\neg$I
2	P	asm, for $\neg$I
⋮	⋮	
4	$\bot$	
5	$\neg P$	$\neg$I 2–4
⋮	⋮	
7	$\bot$	
8	$\neg(P \leftrightarrow \neg P)$	$\neg$I 1–7

Proof. The following is a derivation with no premises and last line $\neg(P \leftrightarrow \neg P)$.

1	$P \leftrightarrow \neg P$	asm, for $\neg$I
2	P	asm, for $\neg$I
3	$\neg P$	$\leftrightarrow$E 1, 2
4	$\bot$	$\neg$E 2, 3
5	$\neg P$	$\neg$I 2–4
6	P	$\leftrightarrow$E 1, 5
7	$\bot$	$\neg$E 6, 5
8	$\neg(P \leftrightarrow \neg P)$	$\neg$I 1–7

Line 5 may be cited at lines 6 and 7: it lies in the box 1–7, as they do. No rule of Section 2.5 is used.

■

In words: no statement is equivalent to its own negation; a definition that made a term mean the failure of that very term would define nothing.

For an instance, let x be a real number. The sentence "$x > 0$ if and only if $x \leq 0$" is false whatever x is: by the convention of Section 2.1, $x \leq 0$ is $\neg(x > 0)$, so the sentence is an instance of $P \leftrightarrow \neg P$, and the lemma, used as a derived rule, refutes every instance.

Quick check. In the proof of Lemma 2.4.2, which two lines use the biconditional, and in which direction?

Solution. Line 3, from the left side P (line 2) to the right side $\neg P$; and line 6, from the right side $\neg P$ (line 5) back to the left side P.

2.4.3 Provable equivalence

Proposition 2.4.1 is a pair of sequents, one in each direction. It is not the first such pair: Proposition 1.4.2 and Proposition 1.6.1 are pairs too. When two statements can each be derived from the other, either may stand in for the other in any derivation, and the situation deserves a name.

Two statements are provably equivalent when each can be derived from the other.

Definition 2.4.2. Provable Equivalence

Statements P and Q are **provably equivalent**, written $P \dashv\vdash Q$, if $P \vdash Q$ **and** $Q \vdash P$.

In words: there is a derivation with premise P and last line Q, and **another** with premise Q and last line P. The symbol is the turnstile of Definition 1.2.3 written twice, back to back; two derivations prove the claim.

A first example is $P \wedge Q \dashv\vdash Q \wedge P$: Proposition 1.2.1 derives $Q \wedge P$ from $P \wedge Q$, and the same derivation with P and Q exchanged derives $P \wedge Q$ from $Q \wedge P$. By the same exchange, Example (First derivations with "or") gives $P \vee Q \dashv\vdash Q \vee P$, and part (a) of Example (An iff read both ways) gives $P \leftrightarrow Q \dashv\vdash Q \leftrightarrow P$.

Second reading · *Edge cases of provable equivalence*

The degenerate example is $P \dashv\vdash P$: the one-line derivation of part (a) of Example (Derivations, simplest first) serves in both directions.

A minimal change gives a one-way street. Example (Two negations from one premise) derived $P \vdash \neg\neg P$, but the converse, $\neg\neg P \vdash P$, would have to remove two negations, and no rule met

so far removes one. So we cannot claim $P \dashv\vdash \neg\neg P$: the clause "$Q \vdash P$" is the one missing. (It holds once Section 2.5 adds its rule.) Likewise Proposition 2.3.2 derives $\neg Q \to \neg P$ from $P \to Q$, but not yet the reverse.

The connective $\leftrightarrow$ and the relation $\dashv\vdash$ are two sides of one idea; the next result makes the link exact, and shows that $\dashv\vdash$ behaves as an equivalence should.

Proposition 2.4.3. Provable Equivalence and the Biconditional

For all statements P, Q, R:

(a) $P \dashv\vdash Q$ if and only if $\vdash P \leftrightarrow Q$;

(b) $P \dashv\vdash P$; if $P \dashv\vdash Q$, then $Q \dashv\vdash P$; and if $P \dashv\vdash Q$ and $Q \dashv\vdash R$, then $P \dashv\vdash R$.

In words: one derivation of $P \leftrightarrow Q$ from no premises is worth two derivations, one each way; and $\dashv\vdash$ can be used like $=$ in a calculation.

Second reading · *The proof: copying derivations*

Idea. Everything here moves whole derivations around. To prove $\vdash P \leftrightarrow Q$, the two boxes of $\leftrightarrow$I are the two given derivations, with their premises turned into the boxes' assumptions. To use $\vdash P \leftrightarrow Q$, write out its derivation and apply $\leftrightarrow$E. For transitivity, run one derivation into the next: the last line of the first is the premise of the second.

Proof. We use one observation twice. Let D be a derivation, and copy its lines, in order and with their numbers shifted, to some place in another derivation, with every line of the copy inside the same boxes as that place. Then every citation inside the copy is still available: if line i of D is available at line k of D, every box of D containing line i contains line k, and the boxes added around the copy contain both lines. Closed boxes cited in D are handled the same way. So each copied line is justified by the same rule as before.

($\Rightarrow$) in (a). Suppose $P \vdash Q$ and $Q \vdash P$, by derivations D_1 and D_2. Build a derivation with no premises as follows. Open a box with the assumption P and, below it, copy D_1, turning each of its premise lines P into a repetition of the assumption by (D3) of Definition 1.4.2. By the observation, the copy is justified, and its last line Q is a line of the box itself, since the last line of D_1 lies in none of D_1's boxes. Close the box. Open a second box with Q, copy D_2 in the same way, and close it on P. Then $\leftrightarrow$I on the two boxes gives $P \leftrightarrow Q$. So $\vdash P \leftrightarrow Q$.

($\Leftarrow$) in (a). Suppose $\vdash P \leftrightarrow Q$, by a derivation D with no premises. For $P \vdash Q$, start with the premise P, copy D below it, and write Q by $\leftrightarrow$E from the last line of the copy, $P \leftrightarrow Q$, and line 1. Both cited lines lie in no box, so both are available. For $Q \vdash P$, do the same with premise Q, using the other form of $\leftrightarrow$E. So $P \dashv\vdash Q$.

(b) The one-line derivation consisting of the premise P shows $P \vdash P$, in both directions, so $P \dashv\vdash P$. Symmetry holds because Definition 2.4.2 says the same thing about P, Q as about Q, P. For transitivity it is enough to show that $P \vdash Q$ and $Q \vdash R$ give $P \vdash R$,

since the reverse direction is the same statement with P and R exchanged. Take a derivation D_1 of Q from P, and below its last line copy a derivation D_2 of R from Q, turning each premise line Q of D_2 into a repetition of the last line of D_1. That line lies in no box, so it is available everywhere below it, and by the observation the rest of the copy is justified. The result has premise P and last line R. This proves the proposition.

■

Here are equivalences derived so far. Each row gives two derived rules (Definition 2.3.1), one in each direction.

Provable equivalence	Derived in
$P \wedge Q \dashv\vdash Q \wedge P$	Proposition 1.2.1, with the letters exchanged for the converse
$P \vee Q \dashv\vdash Q \vee P$	Example (First derivations with "or"), likewise
$(P \wedge Q) \to R \dashv\vdash P \to (Q \to R)$	Proposition 1.4.2
$P \wedge (Q \vee R) \dashv\vdash (P \wedge Q) \vee (P \wedge R)$	Proposition 1.6.1
$(P \to R) \wedge (Q \to R) \dashv\vdash (P \vee Q) \to R$	Example (Cases on a hypothesis are two implications) and Example (A disjunctive hypothesis splits into two implications)
$P \leftrightarrow Q \dashv\vdash (P \to Q) \wedge (Q \to P)$	Proposition 2.4.1
$P \leftrightarrow Q \dashv\vdash Q \leftrightarrow P$	Example (An iff read both ways) (a), with the letters exchanged
$\neg\neg\neg P \dashv\vdash \neg P$	Example (Three negations become one), and Example (Two negations from one premise) (a) with $\neg P$ for P

The fourth row is distributivity of $\wedge$ over $\vee$ as a provable equivalence. Some familiar laws are not in the table, because the text has derived them in one direction only; associativity of $\vee$ is one (Example (Regrouping a disjunction)).

Warning. A sequent in one direction is not an equivalence. $P \wedge Q \vdash P$ holds by $\wedge$E, but the form "from P conclude $P \wedge Q$" breaks: take P: "$2 > 1$" and Q: "$1 > 2$", and the premise is true while the conclusion is false. Before writing $\dashv\vdash$, check that **both** derivations exist.

2.4.4 Proving an iff: two directions

In a written proof, the two boxes of $\leftrightarrow$I become two paragraphs: ($\Rightarrow$) assumes the left side and reaches the right, ($\Leftarrow$) the reverse. Neither may borrow the other's assumption: they are two separate boxes.

The first example completes the story of parity. We use the facts (E1)–(E3) of Section 2.1: every integer is even or odd; no integer is both even and odd; the square of an odd integer is odd.

Theorem 2.4.4. Odd Square, Odd Number

Let n be an integer. Then n is odd if and only if n^2 is odd.

Idea. Two boxes. Forwards is (E3). Backwards, "n^2 is odd" gives nothing to compute with, so split on (E1) and show the even case impossible, as in Example (An impossible case).

Proof. ($\Rightarrow$) This is (E3).

($\Leftarrow$) Suppose n^2 is odd [assumption for the second box]. By (E1), n is even or odd, and we argue by cases [$\vee$E]. If n is odd, we are done. If n is even, say $n = 2k$ with k an integer, then $n^2 = 2(2k^2)$ is even; so n^2 is both even and odd [$\wedge$I], which contradicts (E2) [$\neg$E]. So this case cannot occur, and n is odd in it too [$\bot$E]. In both cases n is odd.

This proves the theorem [$\leftrightarrow$I].

■

With the derived rules of Section 2.3, ($\Leftarrow$) fits in one breath, "n cannot be even, since then n^2 would be even; so n is odd": $\neg$I with (E2), then MT, then DS on (E1).

The next example is about real numbers, and its two directions do different work. Each hypothesis is needed: dropping either one breaks one of the two directions, as the remark after the proof shows.

Theorem 2.4.5. Squaring Preserves Order

Let x and y be real numbers with $x \geq 0$ and $y \geq 0$. Then

$$x \leq y \iff x^2 \leq y^2.$$

Idea. Forwards, multiply $x \leq y$ by $x \geq 0$ and by $y \geq 0$, and chain. Backwards, nothing undoes squaring without square roots, so we split instead: either $x \leq y$, which is the goal, or $y < x$. In the second case the forward argument, run with a strict inequality, gives $y^2 < x^2$, which clashes with the hypothesis: an impossible case, closed by $\bot$E.

Proof. ($\Rightarrow$) Suppose $x \leq y$. Multiplying by $x \geq 0$ gives $x^2 \leq xy$, and multiplying by $y \geq 0$ gives $xy \leq y^2$. Hence $x^2 \leq y^2$.

($\Leftarrow$) Suppose $x^2 \leq y^2$. For real numbers, $x \leq y$ or $y < x$, and we argue by cases [$\vee$E]. If $x \leq y$, we are done. If $y < x$, then $x > y \geq 0$, so $x > 0$. Multiplying $y < x$ by $x > 0$ gives $xy < x^2$, and multiplying $y < x$ by $y \geq 0$ gives $y^2 \leq xy$. Hence $y^2 < x^2$, which contradicts $x^2 \leq y^2$ [$\neg$E]; so this case cannot occur, and $x \leq y$ holds in it [$\bot$E]. In both cases $x \leq y$.

This proves the theorem [$\leftrightarrow$I].

■

Remark. Each hypothesis has a job. The direction ($\Rightarrow$) needs $x \geq 0$, and fails without it: $x = -2$, $y = 1$ gives $x \leq y$ but $x^2 = 4 > 1 = y^2$. (Its proof also multiplied by $y \geq 0$, but there $y \geq 0$ already follows from $y \geq x \geq 0$.) The direction ($\Leftarrow$) used only $y \geq 0$, and fails without it: $x = 1$, $y = -2$ gives $x^2 = 1 \leq 4 = y^2$ but $x > y$.

A payoff comes at once. For $a, b \geq 0$, apply Theorem 2.4.5 to $x = \sqrt{a}$ and $y = \sqrt{b}$: $\sqrt{a} \leq \sqrt{b}$ if and only if $a \leq b$. Taking square roots preserves order, a fact used constantly in estimates. (That square roots of non-negative numbers exist is a school fact here.)

Quick check. Here is a proof of the statement "for a real number x, $x^2 = 1$ if and only if $x = 1$ or $x = -1$".

> If $x = 1$, then $x^2 = 1 \cdot 1 = 1$. If $x = -1$, then $x^2 = (-1)(-1) = 1$. This proves the claim.

Which direction has been proved, and which is missing? Supply the missing one.

Solution. The proof assumes the right side, "$x = 1$ or $x = -1$", splits it into its two cases, and reaches $x^2 = 1$ in both. So it proves ($\Leftarrow$), by $\vee$E. The direction ($\Rightarrow$) is missing.

($\Rightarrow$) Suppose $x^2 = 1$. Then $(x - 1)(x + 1) = x^2 - 1 = 0$. A product of real numbers is zero only if one of the factors is zero, so $x - 1 = 0$ or $x + 1 = 0$, that is, $x = 1$ or $x = -1$.

Warning. A proof of an iff can prove one direction twice and look complete. For "n is odd if and only if n^2 is odd", a proof that shows "if n is odd, then n^2 is odd" and then "if n^2 is not odd, then n is not odd" has proved ($\Rightarrow$) and its contrapositive (Proposition 2.3.2), which is ($\Rightarrow$) again. The direction ($\Leftarrow$), "if n^2 is odd, then n is odd", is still missing; its contrapositive would be "if n is not odd, then n^2 is not odd". Before writing "this proves the theorem", check that one paragraph assumes each side.

2.4.5 What a definition says

Here is a definition written the way this book and most others write definitions.

> A function $f : \mathbb{R} \to \mathbb{R}$ is **strictly increasing** if, for all real numbers x and y, $x < y$ implies $f(x) < f(y)$.

Grammatically the sentence is an implication, "f is strictly increasing **if** [condition]", and read as a theorem it would allow only one move, from the condition to the term. But a definition is not a theorem. It introduces a new term, and the new term means **exactly** the condition: no more, since nothing else was said about the term, and no less, since the condition is all the term was made to express. So the definition is used in both directions, as an $\leftrightarrow$ between the term and the condition, and $\leftrightarrow$E is how it is used.

In a definition, "if" means "if and only if".

This is a convention of mathematical language, not a rule of logic. Some books write "iff" or $:\Leftrightarrow$

in definitions; this book, like most, writes "if" and puts the term defined in **bold**, which signals that the "if" is an "iff". Here it is used three ways.

Example. A definition used in both directions

(a) Show that $f(x) = 3x - 5$ defines a strictly increasing function $f\colon \mathbb{R} \to \mathbb{R}$.

(b) Suppose $g\colon \mathbb{R} \to \mathbb{R}$ is strictly increasing. Show that $g(1) < g(2)$.

(c) Show that $h(x) = x^2$ does not define a strictly increasing function $h\colon \mathbb{R} \to \mathbb{R}$.

Solution.

(a) Let x and y be real numbers with $x < y$. Multiplying by $3 > 0$ gives $3x < 3y$, and subtracting 5 gives $3x - 5 < 3y - 5$, that is, $f(x) < f(y)$. So f satisfies the condition, and by the definition, read from right to left, f is strictly increasing [$\leftrightarrow$E].

(b) Since g is strictly increasing, the definition, read from left to right, says that $x < y$ implies $g(x) < g(y)$ for all real x and y [$\leftrightarrow$E]. Taking $x = 1$ and $y = 2$, and using $1 < 2$, we get $g(1) < g(2)$ [$\to$E].

(c) Suppose h were strictly increasing [assumption for $\neg$I]. By the definition read from left to right, as in (b), $-1 < 0$ would give $h(-1) < h(0)$, that is, $1 < 0$, which is false [$\neg$E]. So h is not strictly increasing [$\neg$I].

Part (a) uses the definition from the condition to the term; parts (b) and (c) use it from the term to the condition. A definition read only as an "if" would permit (a) and nothing else. Because the term means exactly the condition, a candidate condition is judged against the intended meaning. Had the definition said "$x < y$ implies $f(x) \leq f(y)$", it would be too broad: the constant function $f(x) = 0$ would count as strictly increasing. Had it said "$x < y$ implies $f(y) - f(x) \geq 1$", it would be too narrow: $f(x) = x$ would not count, since $0 < \frac{1}{2}$ and $\frac{1}{2} - 0 < 1$.

Warning. The convention is for **definitions only**. In a theorem, "if" means "if", one direction. The theorem "if $6 \mid n$, then n is even" is true for every integer n: $n = 6k = 2(3k)$. Its converse, "if n is even, then $6 \mid n$", breaks at $n = 4$. Reading a theorem's "if" as "iff" means proving its converse without a proof, which is the error of Argument B in Section 1.1.

↻ **Second reading** · *The same slip in calculus*

A familiar theorem of calculus says: if f is differentiable at a, then f is continuous at a. Its converse is false. The function $f(x) = |x|$ is continuous at 0, but it is not differentiable there, since its graph has a corner: the quotient $|h|/h$ is 1 for $h > 0$ and -1 for $h < 0$.

Quick check. Write the converse of "if a real number x satisfies $x > 2$, then $x^2 > 4$", and break it.

Solution. The converse is "if $x^2 > 4$, then $x > 2$". The value $x = -3$ breaks it: $(-3)^2 = 9 > 4$, and $-3 > 2$ fails.

2.4.6 Chains of equivalences

"The following are equivalent: (a), (b), (c)" means that any two of the three are equivalent: three iff statements, or six implications. Three implications suffice, if they go round in a cycle:

$$\text{(a)} \Rightarrow \text{(b)} \Rightarrow \text{(c)} \Rightarrow \text{(a)}.$$

Once the cycle is proved, any condition leads to any other by following the arrows round. Here is the derivation behind that sentence.

Proposition 2.4.6. Proving Equivalence by a Cycle

$P \to Q,\ Q \to R,\ R \to P \vdash P \leftrightarrow Q$. Likewise, from the same three premises, $Q \leftrightarrow R$ and $R \leftrightarrow P$ are derivable.

Idea. The goal is an iff, so there are two boxes. From P, one step along the cycle reaches Q. From Q, the cycle reaches P the long way round, through R: a backward chain of $\to$E, as in Section 1.3. The other two equivalences need no new work, because the three premises do not change when the letters are rotated.

The goal at the foot (left), then its two boxes (right).

1	$P \to Q$
2	$Q \to R$
3	$R \to P$
	⋮
9	$P \leftrightarrow Q$

1	$P \to Q$	
2	$Q \to R$	
3	$R \to P$	
4	$\quad P$	asm, for $\leftrightarrow$I
	$\quad \vdots$	
5	$\quad Q$	
6	$\quad Q$	asm, for $\leftrightarrow$I
	$\quad \vdots$	
8	$\quad P$	
9	$P \leftrightarrow Q$	$\leftrightarrow$I 4–5, 6–8

Proof. The following is a derivation with the three premises and last line $P \leftrightarrow Q$.

1	$P \to Q$	
2	$Q \to R$	
3	$R \to P$	
4	$\quad P$	asm, for $\leftrightarrow$I
5	$\quad Q$	$\to$E 1, 4
6	$\quad Q$	asm, for $\leftrightarrow$I
7	$\quad R$	$\to$E 2, 6
8	$\quad P$	$\to$E 3, 7
9	$P \leftrightarrow Q$	$\leftrightarrow$I 4–5, 6–8

Now replace P, Q, R throughout by Q, R, P respectively. The premises $P \to Q$, $Q \to R$, $R \to P$ become $Q \to R$, $R \to P$, $P \to Q$, the same three statements in a different order, and every line remains justified by the same rule, since the letters stood for any statements (as in clause 2 after Definition 2.3.1). So the renamed derivation derives $Q \leftrightarrow R$ from the same premises. Rotating once more gives $R \leftrightarrow P$. This proves the proposition.

■

In a cycle every condition reaches every other along the arrows; longer cycles work the same way, with longer chains of $\to$E in the boxes.

Warning. Two arrows back and forth are not a cycle through three conditions. Proving (a) $\Rightarrow$ (b) and (b) $\Rightarrow$ (a) shows that (a) and (b) are equivalent and says nothing about (c): no arrow enters or leaves it. For a real number x, take (a) "$x > 0$", (b) "$2x > 0$", and (c) "$x > 1$". Both arrows between (a) and (b) are true, and (c) is not equivalent to them: $x = \frac{1}{2}$ satisfies (a) and (b) but not (c). Every condition in the list must lie on the cycle.

Example. Three descriptions of the unit interval

Let x be a real number. Prove that the following are equivalent.

(a) $0 < x < 1$.
(b) $x^2 < x$.
(c) $x > 0$ and $1/x > 1$.

Solution. By Proposition 2.4.6 it suffices to prove (a) $\Rightarrow$ (b) $\Rightarrow$ (c) $\Rightarrow$ (a).

(a) $\Rightarrow$ (b). Suppose $0 < x < 1$. Multiplying $x < 1$ by $x > 0$ gives $x^2 < x$.

(b) $\Rightarrow$ (c). Suppose $x^2 < x$. First, $x > 0$: otherwise $x \le 0$, and then $x \le 0 \le x^2$, so $x \le x^2$, which contradicts $x^2 < x$. (In rules: we have $x \le 0$ or $x > 0$; the first is refuted by a $\neg$I box, and DS (Proposition 2.3.4) leaves the second.) Since $x > 0$, we may divide $x^2 < x$ by x, which gives $x < 1$, and then divide $x < 1$ by x, which gives $1 < 1/x$. So $x > 0$ and $1/x > 1$.

(c) $\Rightarrow$ (a). Suppose $x > 0$ and $1/x > 1$. Multiplying $1 < 1/x$ by $x > 0$ gives $x < 1$. So

$0 < x < 1$.

The order of the cycle was a choice, made so that each arrow was short: (b) $\Rightarrow$ (c) carried the only case split. The reverse cycle would also do, with different arrows to prove.

2.4.7 The rule card

Before the card, the two rules at work on large statements.

Example. Instances of $\leftrightarrow$I and $\leftrightarrow$E

Rule	Letters replaced	Cited	New line
$\leftrightarrow$I	$\varphi := P \wedge Q$, $\psi := Q \wedge P$	box $P \wedge Q$ … $Q \wedge P$, box $Q \wedge P$ … $P \wedge Q$	$(P \wedge Q) \leftrightarrow (Q \wedge P)$
$\leftrightarrow$E	$\varphi := \neg P$, $\psi := Q \to R$	$\neg P \leftrightarrow (Q \to R)$, $Q \to R$	$\neg P$ (the second form)
$\leftrightarrow$E	$\varphi := P \vee Q$, $\psi := \neg R$	$(P \vee Q) \leftrightarrow \neg R$, $P \vee Q$	$\neg R$

Not an instance. From two lines $P \to Q$ and $Q \to P$, the line "$P \leftrightarrow Q$ — $\leftrightarrow$I" is not an instance: $\leftrightarrow$I cites two boxes, not two implications. The two lines give the iff through Proposition 2.4.1 used as a derived rule (after $\wedge$I), or by opening the two boxes.

Chapter 1's seven rules are on the card at the end of Section 1.6; here are this chapter's so far.

Rule	Shape	You write	Lean
$\neg$E	$\varphi, \neg\varphi \Longrightarrow \bot$	"…, which contradicts …"	`exact hn hp` (goal `False`)
$\bot$E	$\bot \Longrightarrow$ any φ	"This case cannot occur." "So there is nothing to prove."	`exfalso`
$\neg$I	box φ … $\bot \Longrightarrow \neg\varphi$	"Suppose P. … This is a contradiction. Hence not P."	`intro hp` (goal `False`)
MT	$\varphi \to \psi, \neg\psi \Longrightarrow \neg\varphi$ (derived)	"Since Q fails, so does P."	`mt h hnq`
CP	$\varphi \to \psi \Longrightarrow \neg\psi \to \neg\varphi$ (derived)	"We prove the contrapositive: suppose not Q; … so not P."	`mt`
DS	$\varphi \vee \psi, \neg\varphi \Longrightarrow \psi$ (derived)	"P is impossible, so Q."	`h.resolve_left hnp`
$\leftrightarrow$I	box φ … ψ, box ψ … $\varphi \Longrightarrow \varphi \leftrightarrow \psi$	"($\Rightarrow$) Suppose P. … so Q. ($\Leftarrow$) Suppose Q. … so P."	`constructor`

Rule	Shape	You write	Lean
$\leftrightarrow$E	$\varphi \leftrightarrow \psi$, $\varphi \Longrightarrow \psi$ (or with ψ, φ)	"Since P holds if and only if Q does, and P holds, so does Q."	`h.mp hp` / `h.mpr hq`

The rows for this section's rules complete the table for the connectives. The Lean names in the last column come from the two fields of an iff, `mp` ("modus ponens", left to right) and `mpr` (right to left).

Warning (Not allowed: $\leftrightarrow$I and $\leftrightarrow$E). Each step below is wrong; the tag names the clause it breaks.

- Close $\leftrightarrow$I with one box: clause 1, two boxes (a single box proves $\varphi \to \psi$ only).
- In the second box, cite a line of the first: clause 2, the boxes are siblings (the pair below).
- Write $\neg\psi$ by $\leftrightarrow$E from $\varphi \leftrightarrow \psi$ and $\neg\varphi$: clause 4, $\leftrightarrow$E needs one side itself; $\neg\psi$ takes a $\neg$I box.
- Write $\varphi \leftrightarrow \psi$ by $\leftrightarrow$I citing two implication lines: clause 3, the rule cites boxes, as ranges (the look-alike of Example (Instances of $\leftrightarrow$I and $\leftrightarrow$E)).

Clause 2 is where the rule bites: WRONG beside RIGHT, with premises $P \to Q$, R and $(Q \wedge R) \to P$. On the left, line 7 copies P from the closed first box.

1	$P \to Q$	
2	R	
3	$(Q \wedge R) \to P$	
4	$\quad P$	asm, for $\leftrightarrow$I
5	$\quad Q$	$\to$E 1, 4
6	$\quad Q$	asm, for $\leftrightarrow$I
7	$\quad P$	R 4
8	$P \leftrightarrow Q$	$\leftrightarrow$I 4–5, 6–7

1	$P \to Q$	
2	R	
3	$(Q \wedge R) \to P$	
4	$\quad P$	asm, for $\leftrightarrow$I
5	$\quad Q$	$\to$E 1, 4
6	$\quad Q$	asm, for $\leftrightarrow$I
7	$\quad Q \wedge R$	$\wedge$I 6, 2
8	$\quad P$	$\to$E 3, 7
9	$P \leftrightarrow Q$	$\leftrightarrow$I 4–5, 6–8

The left list never cites premises 2 and 3, so it would show $P \to Q \vdash P \leftrightarrow Q$, which breaks: with P: "$x > 5$" and Q: "$x > 3$", at $x = 4$ the premise is true and $P \leftrightarrow Q$ is false.

Remark (Looking ahead). Chapter 5 uses the lemma $\neg(P \leftrightarrow \neg P)$ to show that no barber can shave exactly those who do not shave themselves, and in what sense no set can contain exactly the sets that do not contain themselves. *Analysis on the Real Line*, Chapter 1 proves from its axioms that every non-negative real number has a square root, the school fact behind the square-root payoff above.

2.4.8 Exercises

2.4.8.1 A. Check your understanding

Exercise. A1

(a) State the rules $\leftrightarrow$I and $\leftrightarrow$E, and the definition of $P \dashv\vdash Q$.

(b) For a real number $x \geq 0$, consider "$x^2 = x$ if and only if $x = 0$ or $x = 1$". Write the directions ($\Rightarrow$) and ($\Leftarrow$) as implications. Which one would a proof begin with "Suppose $x^2 = x$"?

(c) True or false, with a reason: from the definition "an integer n is **even** if there is an integer k with $n = 2k$" and the fact that 10 is even, we may conclude that $10 = 2k$ for some integer k.

(d) True or false, with a reason: from the theorem "if $x > 1$, then $x^2 > 1$" and the fact that $x^2 > 1$, we may conclude that $x > 1$.

Solution.

(a) $\leftrightarrow$I: from a box $\varphi \dots \psi$ and a box $\psi \dots \varphi$, both available, write $\varphi \leftrightarrow \psi$. $\leftrightarrow$E: from $\varphi \leftrightarrow \psi$ and either side, write the other (Definition 2.4.1). $P \dashv\vdash Q$ means $P \vdash Q$ and $Q \vdash P$ (Definition 2.4.2).

(b) ($\Rightarrow$): if $x^2 = x$, then $x = 0$ or $x = 1$. ($\Leftarrow$): if $x = 0$ or $x = 1$, then $x^2 = x$. The direction ($\Rightarrow$) begins "Suppose $x^2 = x$", since it assumes the left side.

(c) True. A definition's "if" means "if and only if", so the definition may be used from the term "even" to the condition, by $\leftrightarrow$E. (Indeed $k = 5$.)

(d) False. A theorem's "if" is one direction only, and the conclusion would be its converse. The converse fails at $x = -2$: $x^2 = 4 > 1$ but $x < 1$.

Exercise. A2: Drill

The parts run from one step, through a missing line, a justification column and steps to judge, to short derivations.

(a) Lines 1 $P \leftrightarrow Q$, 2 Q. What does "$\leftrightarrow$E 1, 2" write?

(b) A box 3–6 runs from $P \vee Q$ to $\neg R$, and a box 7–9 from $\neg R$ to $P \vee Q$. What does "$\leftrightarrow$I 3–6, 7–9" write?

(c) Write "$x^2 = 4$ if and only if $x = 2$ or $x = -2$" as two implications. Which begins "Suppose $x^2 = 4$"?

(d) Definition: "an integer n is a **multiple of 3** if $n = 3k$ for some integer k". From $12 = 3 \cdot 4$, what follows, reading the definition which way?

(e) Line 6 is $Q \wedge R$, justified "$\leftrightarrow$E 2, 5", and line 2 is $P \leftrightarrow (Q \wedge R)$. What is line 5?

(f) Line 12 is $\neg P \leftrightarrow Q$, justified "$\leftrightarrow$I 4–7, 8–11". What are lines 4, 7, 8 and 11?

(g) Fill in the justifications (i)–(iv).

1	$P \leftrightarrow \neg Q$	
2	Q	
3	$\quad P$	(i)
4	$\quad \neg Q$	(ii)
5	$\quad \bot$	(iii)
6	$\neg P$	(iv)

(h) Lines 1 $P \leftrightarrow Q$, 2 $\neg P$. Does $\leftrightarrow$E give anything?

(i) A box 2–4 runs from P to Q; a box 5–7 from Q to P, and line 6 cites line 3. Is "$P \leftrightarrow Q$ — $\leftrightarrow$I 2–4, 5–7" correct?

(j) One box runs from P to Q. What does it prove?

(k) True or false: $P \dashv\vdash Q$ says no more than $\vdash P \to Q$.

(l) A proof that (a), (b), (c) are equivalent proves (a) $\Rightarrow$ (b), (b) $\Rightarrow$ (a) and (b) $\Rightarrow$ (c). Is it complete?

(m) Derive $P \leftrightarrow Q,\ P \vdash Q \wedge P$.

(n) Derive $P \leftrightarrow Q,\ Q \leftrightarrow R,\ R \vdash P$.

(o) Derive $P \leftrightarrow Q,\ \neg P \vdash \neg Q$.

Solution.

(a) P, by the second form.

(b) $(P \vee Q) \leftrightarrow \neg R$: the first box's assumption on the left, its last line on the right.

(c) ($\Rightarrow$) if $x^2 = 4$, then $x = 2$ or $x = -2$; ($\Leftarrow$) the converse. The direction ($\Rightarrow$) begins "Suppose $x^2 = 4$".

(d) 12 is a multiple of 3, reading the definition from the condition to the term ($\leftrightarrow$E).

(e) P, the left side, so that the first form writes the right side $Q \wedge R$.

(f) Line 4 is $\neg P$ and line 7 is Q (the first box); line 8 is Q and line 11 is $\neg P$ (the second).

(g) (i) assumption, for $\neg$I; (ii) $\leftrightarrow$E 1, 3; (iii) $\neg$E 2, 4; (iv) $\neg$I 3–5.

(h) No: $\leftrightarrow$E needs P or Q itself. $\neg Q$ comes by a $\neg$I box, as in part (o).

(i) No: line 6 cites line 3, inside the closed sibling box 2–4.

(j) $P \to Q$, by $\to$I; not $P \leftrightarrow Q$.

(k) False: it says $P \vdash Q$ and $Q \vdash P$, two derivations (equivalently $\vdash P \leftrightarrow Q$); $\vdash P \to Q$ is only half of it.

(l) No: no arrow leaves (c). Add (c) $\Rightarrow$ (a) or (c) $\Rightarrow$ (b).

(m) 3 Q, $\leftrightarrow$E 1, 2; 4 $Q \wedge P$, $\wedge$I 3, 2.

(n) 4 Q, $\leftrightarrow$E 2, 3; 5 P, $\leftrightarrow$E 1, 4.

(o) 3 Q, assumption, for $\neg$I; 4 P, $\leftrightarrow$E 1, 3; 5 $\bot$, $\neg$E 4, 2; 6 $\neg Q$, $\neg$I 3–5.

2.4.8.2 B. Practice

Exercise. B1: Iff is transitive

Show that $P \leftrightarrow Q,\ Q \leftrightarrow R \vdash P \leftrightarrow R$.

Solution. The goal is an iff, so we open two boxes: P to R, and R to P. In each, the premises are used as bridges, one after the other, by $\leftrightarrow$E.

1	$P \leftrightarrow Q$	
2	$Q \leftrightarrow R$	
3	$\quad P$	asm, for $\leftrightarrow$I
4	$\quad Q$	$\leftrightarrow$E 1, 3
5	$\quad R$	$\leftrightarrow$E 2, 4
6	$\quad R$	asm, for $\leftrightarrow$I
7	$\quad Q$	$\leftrightarrow$E 2, 6
8	$\quad P$	$\leftrightarrow$E 1, 7
9	$P \leftrightarrow R$	$\leftrightarrow$I 3–5, 6–8

Lines 4 and 8 use $P \leftrightarrow Q$ in opposite directions, as do lines 5 and 7 with $Q \leftrightarrow R$.

Exercise. B2: A parity equivalence

Let n be an integer. Prove that n is even if and only if $n+3$ is odd.

Solution. ($\Rightarrow$) Suppose n is even, say $n = 2k$ with k an integer. Then $n+3 = 2k+2+1 = 2(k+1)+1$, and $k+1$ is an integer, so $n+3$ is odd.

($\Leftarrow$) Suppose $n+3$ is odd, say $n+3 = 2j+1$ with j an integer. Then $n = 2j-2 = 2(j-1)$, and $j-1$ is an integer, so n is even.

This proves the equivalence.

Exercise. B3: Strict inequalities

Let x and y be real numbers with $x \geq 0$ and $y \geq 0$. Deduce from Theorem 2.4.5 and Example (Negating both sides of an iff) that $x < y$ if and only if $x^2 < y^2$. You may use that, for real numbers a and b, $a < b$ holds exactly when $b \leq a$ fails.

(a) Apply Theorem 2.4.5 with the roles of x and y exchanged. What does it say?

(b) Use Example (Negating both sides of an iff) as a derived rule on (a): with which statements in place of P and Q, and what does it give?

(c) Translate both sides of (b) with the fact quoted.

(d) What have you proved?

Solution.

(a) The exchange is allowed, since both numbers are non-negative: $y \leq x$ if and only if $y^2 \leq x^2$.

(b) With P: "$y \leq x$" and Q: "$y^2 \leq x^2$", Example (Negating both sides of an iff), used as a derived rule (Definition 2.3.1), gives: $y \leq x$ fails if and only if $y^2 \leq x^2$ fails.

(c) By the fact quoted, the left side says $x < y$ and the right side says $x^2 < y^2$.

(d) That $x < y$ if and only if $x^2 < y^2$.

Exercise. B4: Both directions

A search for $P \leftrightarrow Q \vdash (P \wedge R) \leftrightarrow (Q \wedge R)$ has reached the stage below. *Name the move, then complete it.*

1	$P \leftrightarrow Q$	
	⋮	
12	$(P \wedge R) \leftrightarrow (Q \wedge R)$	?

Solution. The move is *two boxes for an iff*: the goal's main connective is $\leftrightarrow$, so line 12 will be $\leftrightarrow$I, with a box from $P \wedge R$ to $Q \wedge R$ and a box back. Inside each, take the conjunction apart, cross with $\leftrightarrow$E, and rebuild.

1	$P \leftrightarrow Q$	
2	$\quad P \wedge R$	asm, for $\leftrightarrow$I
3	$\quad P$	$\wedge$E 2
4	$\quad R$	$\wedge$E 2
5	$\quad Q$	$\leftrightarrow$E 1, 3
6	$\quad Q \wedge R$	$\wedge$I 5, 4
7	$\quad Q \wedge R$	asm, for $\leftrightarrow$I
8	$\quad Q$	$\wedge$E 7
9	$\quad R$	$\wedge$E 7
10	$\quad P$	$\leftrightarrow$E 1, 8
11	$\quad P \wedge R$	$\wedge$I 10, 9
12	$(P \wedge R) \leftrightarrow (Q \wedge R)$	$\leftrightarrow$I 2–6, 7–11

The second box never cites a line of the first: line 10 gets P afresh from line 8.

Exercise. B5: Criticize the definition

For real numbers a, b, x, we want to define "x lies **strictly between** a and b", meaning that one of a, b is less than x and x is less than the other. Here are five candidate conditions.

(i) $a < x$ and $x < b$.

(ii) $a < x$ or $x < b$.

(iii) ($a < x$ and $x < b$) or ($b < x$ and $x < a$).

(iv) $(x - a)(x - b) < 0$.

(v) $x \leq a$ fails and $b \leq x$ fails.

(a) For each, say whether it is too broad (it admits a triple a, b, x that should be excluded), too narrow (it excludes one that should be admitted), or right, giving a triple for each fault.

(b) A definition is read as an iff. Prove that (iii) and (iv) are equivalent.

(c) Which candidates would make "if x lies strictly between a and b, then x lies strictly between b and a" false? Give one triple that shows it for all of them.

Solution.

(a) (i) Too narrow: $a = 5$, $b = 2$, $x = 3$ lies between and fails (i). It is not too broad: (i) gives one of the wanted orders.

(ii) Too broad: $a = 2$, $b = 5$, $x = 100$ satisfies $2 < 100$ but does not lie between. And too narrow: $a = 5$, $b = 2$, $x = 3$ lies between, but $5 < 3$ and $3 < 2$ both fail (the triple of (c)).

(iii) Right: it lists the two possible orders.

(iv) Right, by (b).

(v) This is (i) in negative dress: by the convention of Section 2.1, "$x \leq a$ fails" is $a < x$ and "$b \leq x$ fails" is $x < b$. So it is too narrow, by the same triple as (i), and needlessly negative.

(b) ($\Rightarrow$) In the case $a < x < b$, $x - a > 0$ and $x - b < 0$, so the product is negative. In the case $b < x < a$, $x - a < 0$ and $x - b > 0$, so again the product is negative. The two cases are $\vee$E on (iii).
($\Leftarrow$) Suppose the product is negative. Then neither factor is 0, and the two factors cannot have the same sign, since a product of two positive or of two negative numbers is positive. So $x - a > 0 > x - b$, that is $a < x < b$, or $x - a < 0 < x - b$, that is $b < x < a$.

(c) (i), (ii) and (v). With $a = 2$, $b = 5$, $x = 3$, each holds for the ends $2, 5$ and fails for the ends $5, 2$ (for (ii): $2 < 3$, while neither $5 < 3$ nor $3 < 2$). A definition of "between" should not depend on the order of the two ends; (iii) and (iv) do not, since exchanging a and b exchanges the two disjuncts of (iii) and the two factors of (iv).

Marking scheme. (10 points) - 4: (a), one point each for (i) and (v) with a correct triple, one for (ii) with **both** triples (too broad and too narrow), and one for calling (iii) and (iv) right. - 4: (b), each direction 2; the second must say why the factors have opposite signs. - 2: (c), (i), (ii) and (v) with one triple.

Fatal error (0 for (b)): proving only one direction and calling the conditions equivalent.

Exercise. B6: Four conditions, one cycle

Let n be an integer. Prove by one cycle of four implications that the following are equivalent.

(i) n is odd.
(ii) $n+1$ is even.
(iii) $8 \mid n^2 - 1$.
(iv) n^2 is odd.

You may use Theorem 2.4.4 and the fact that $k(k-1)$ is even for every integer k. Then: a proof that proved only (i) $\Leftrightarrow$ (ii) and (iii) $\Leftrightarrow$ (iv) would be incomplete; which two implications would complete it most cheaply?

Solution. We prove the cycle (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (iv) $\Rightarrow$ (i). It suffices by the argument of Proposition 2.4.6: for any two conditions, each reaches the other by a chain of $\to$E along the arrows, inside the two boxes of $\leftrightarrow$I.

(i) $\Rightarrow$ (ii). If $n = 2k+1$, then $n+1 = 2(k+1)$.

(ii) $\Rightarrow$ (iii). If $n+1 = 2j$, then $n-1 = 2(j-1)$, so $n^2 - 1 = (n-1)(n+1) = 4j(j-1)$. By the fact quoted, $j(j-1) = 2m$ for an integer m, so $n^2 - 1 = 8m$.

(iii) $\Rightarrow$ (iv). If $n^2 - 1 = 8m$, then $n^2 = 2(4m) + 1$, which is odd.

(iv) $\Rightarrow$ (i). This is Theorem 2.4.4, read from right to left.

(Exercise B3 of Section 1.6 met the fact (i) $\Rightarrow$ (iii) by another route.)

For the second question: (i) $\Rightarrow$ (iv) and (iv) $\Rightarrow$ (i). Both are Theorem 2.4.4, so they cost nothing. Any two implications joining the two pairs, one in each direction, would do, for instance (ii) $\Rightarrow$ (iii) and (iv) $\Rightarrow$ (i); with them every condition reaches every other.

Marking scheme. (10 points) - 1: a cycle through all four conditions, citing the argument of Proposition 2.4.6 or arguing that it suffices. - 8: 2 for each of the four arrows. - 1: the completion question.

Fatal errors (0 for the step): a "cycle" that omits a condition; (iii) $\Rightarrow$ (iv) or (iv) $\Rightarrow$ (i) argued by examples.

Exercise. B7: Derive or break

For each claim, derive it with the rules of Chapter 1 and Sections 2.1–2.4, or break it by statements about a real number x that make its counterexample set true.

(a) $P \leftrightarrow Q \vdash (P \lor R) \leftrightarrow (Q \lor R)$.
(b) $(P \lor R) \leftrightarrow (Q \lor R) \vdash P \leftrightarrow Q$.
(c) $P \leftrightarrow (Q \land R) \vdash P \to Q$.
(d) $P \leftrightarrow (Q \land R) \vdash Q \to P$.
(e) $P \leftrightarrow Q,\ Q \leftrightarrow \neg R \vdash \neg(P \land R)$.
(f) $P \leftrightarrow Q \vdash \neg(P \land \neg Q)$.

(g) $\neg(P \wedge \neg Q) \vdash P \leftrightarrow Q$.
(h) $P \leftrightarrow Q,\ P \vee Q \vdash P \wedge Q$.

Solution.

(a) Derive. Two boxes for the iff, and a case split inside each.

1	$P \leftrightarrow Q$	
2	$\quad P \vee R$	asm, for $\leftrightarrow$I
3	$\quad\quad P$	asm, for $\vee$E
4	$\quad\quad Q$	$\leftrightarrow$E 1, 3
5	$\quad\quad Q \vee R$	$\vee$I 4
6	$\quad\quad R$	asm, for $\vee$E
7	$\quad\quad Q \vee R$	$\vee$I 6
8	$\quad Q \vee R$	$\vee$E 2, 3–5, 6–7
9	$\quad Q \vee R$	asm, for $\leftrightarrow$I
10	$\quad\quad Q$	asm, for $\vee$E
11	$\quad\quad P$	$\leftrightarrow$E 1, 10
12	$\quad\quad P \vee R$	$\vee$I 11
13	$\quad\quad R$	asm, for $\vee$E
14	$\quad\quad P \vee R$	$\vee$I 13
15	$\quad P \vee R$	$\vee$E 9, 10–12, 13–14
16	$(P \vee R) \leftrightarrow (Q \vee R)$	$\leftrightarrow$I 2–8, 9–15

(b) Break. Take P: "$x > 0$", Q: "$x > 5$", R: "$x > -1$". At $x = 1$, $P \vee R$ and $Q \vee R$ are both true, so the premise is true; P is true and Q false, so $P \leftrightarrow Q$ is false. The counterexample set (Definition 1.1.3) $(P \vee R) \leftrightarrow (Q \vee R)$, $\neg(P \leftrightarrow Q)$ is true at $x = 1$.

(c) Derive.

1	$P \leftrightarrow (Q \wedge R)$	
2	$\quad P$	asm, for $\rightarrow$I
3	$\quad Q \wedge R$	$\leftrightarrow$E 1, 2
4	$\quad Q$	$\wedge$E 3
5	$P \rightarrow Q$	$\rightarrow$I 2–4

(d) Break. Take P: "$x > 5$", Q: "$x > 0$", R: "$x > 3$". At $x = 1$, $Q \wedge R$ and P are both false, so the premise is true, while $Q \rightarrow P$ is false, since Q is true and P false. The counterexample set is true at $x = 1$.

(e) Derive.

1	$P \leftrightarrow Q$	
2	$Q \leftrightarrow \neg R$	
3	$\quad P \wedge R$	asm, for $\neg$I
4	$\quad P$	$\wedge$E 3
5	$\quad Q$	$\leftrightarrow$E 1, 4
6	$\quad \neg R$	$\leftrightarrow$E 2, 5
7	$\quad R$	$\wedge$E 3
8	$\quad \bot$	$\neg$E 7, 6
9	$\neg(P \wedge R)$	$\neg$I 3–8

(f) Derive.

1	$P \leftrightarrow Q$	
2	$\quad P \wedge \neg Q$	asm, for $\neg$I
3	$\quad P$	$\wedge$E 2
4	$\quad Q$	$\leftrightarrow$E 1, 3
5	$\quad \neg Q$	$\wedge$E 2
6	$\quad \bot$	$\neg$E 4, 5
7	$\neg(P \wedge \neg Q)$	$\neg$I 2–6

(g) Break. Take P: "$x > 5$", Q: "$x > 0$". At $x = 1$, $P \wedge \neg Q$ is false, since P is, so the premise is true; P is false and Q true, so $P \leftrightarrow Q$ is false. The counterexample set is true at $x = 1$.

(h) Derive. Split on the second premise; in each case $\leftrightarrow$E supplies the other letter.

1	$P \leftrightarrow Q$	
2	$P \vee Q$	
3	$\quad P$	asm, for $\vee$E
4	$\quad Q$	$\leftrightarrow$E 1, 3
5	$\quad P \wedge Q$	$\wedge$I 3, 4
6	$\quad Q$	asm, for $\vee$E
7	$\quad P$	$\leftrightarrow$E 1, 6
8	$\quad P \wedge Q$	$\wedge$I 7, 6
9	$P \wedge Q$	$\vee$E 2, 3–5, 6–8

(b) and (g) are the converses of (a) and (f), and (d) replaces the conclusion of (c) by its converse; each of the three breaks. (a) is the $\vee$-twin of Exercise (B4: Both directions), with a case split inside each box.

Marking scheme. (16 points) For each claim: 1 for the verdict; 1 for a derivation every line of which checks, or for a choice of statements and one value of x at which every member of

the counterexample set is true. Fatal errors (0 for the claim): a "break" at which a premise fails; RAA in a derivation; a box of $\leftrightarrow$I citing its sibling.

2.4.8.3 C. Going deeper

Exercise. C1: Divisibility by six

For integers d and n, write $d \mid n$ if $n = dk$ for some integer k. Let n be an integer. Prove that the following three conditions are equivalent, by a single cycle of three implications.

(i) $6 \mid n$;
(ii) $2 \mid n$ and $3 \mid n$;
(iii) $6 \mid n^2$.

You may use Theorem 2.1.3 and part (a) of Exercise (B3: The square root of three, used later).

Solution (Hint 1). Choose the order of the cycle so that each arrow is short. The two results you may use both pass from a square to the number itself.

Solution (Hint 2). Try (i) $\Rightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i). The last arrow is the one that needs an idea: the two halves of (ii) must be used together, not one at a time.

Solution. We prove (i) $\Rightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i), which suffices by Proposition 2.4.6.

(i) $\Rightarrow$ (iii). Suppose $n = 6k$. Then $n^2 = 36k^2 = 6(6k^2)$, so $6 \mid n^2$.

(iii) $\Rightarrow$ (ii). Suppose $n^2 = 6j$. Then $n^2 = 2(3j)$ is even, so n is even by Theorem 2.1.3, that is, $2 \mid n$. Also $n^2 = 3(2j)$, so $3 \mid n^2$, and by part (a) of Exercise (B3: The square root of three, used later), $3 \mid n$. Hence $2 \mid n$ and $3 \mid n$ [$\wedge$I].

(ii) $\Rightarrow$ (i). Suppose $2 \mid n$ and $3 \mid n$, and write $n = 3b$ with b an integer. The integer b is odd or even. If b were odd, then $n = 3b$ would be odd: writing $b = 2c + 1$ gives $3b = 6c + 3 = 2(3c + 1) + 1$. But n is even, and no integer is both, so this case cannot occur [$\neg$E, $\bot$E]. Hence b is even, say $b = 2c$, and $n = 6c$, so $6 \mid n$.

Marking scheme. (10 points) - 1: names a cycle through all three conditions and cites Proposition 2.4.6 (or proves the needed implications explicitly). - 2: (i) $\Rightarrow$ (iii) by computation. - 3: (iii) $\Rightarrow$ (ii), citing Theorem 2.1.3 for the factor 2 and part (a) of Exercise (B3: The square root of three, used later) for the factor 3. - 4: (ii) $\Rightarrow$ (i), combining the two factors correctly (ruling out b odd, or $n = 3n - 2n$, or otherwise).

Fatal errors (0 for the step): a "cycle" that omits a condition; using "$2 \mid n$ and $3 \mid n$ gives $6 \mid n$" without proof; using "$6 \mid n^2$ implies $6 \mid n$" as known. Any other cycle through the three conditions earns full marks if each of its arrows is proved: 1 point for the cycle and its citation, and 3 points for each of its three arrows.

Exercise. C2: Dropping a hypothesis, and repairing the statement

(a) Suppose only $y \geq 0$ is assumed in Theorem 2.4.5, and x is an arbitrary real number. The remark after it says that ($\Leftarrow$) survives. Which steps of its proof use $y \geq 0$, and why is $x \geq 0$ never needed? Refute ($\Rightarrow$) with a new pair.

(b) Prove that for **all** real numbers x and y, $|x| \leq |y|$ if and only if $x^2 \leq y^2$.

Hint: for (b), apply Theorem 2.4.5 to the right pair of numbers.

Solution.

(a) The proof of ($\Leftarrow$) uses $y \geq 0$ twice: for $x > 0$ from $x > y$, and for $y^2 \leq xy$. Its other multiplication is by the derived $x > 0$, so $x \geq 0$ is never used. For ($\Rightarrow$), take $x = -3$, $y = 1$: $x \leq y$ and $y \geq 0$, but $x^2 = 9 > 1 = y^2$.

(b) Let x and y be real numbers. Then $|x| \geq 0$ and $|y| \geq 0$, so Theorem 2.4.5 applies to the pair $|x|$, $|y|$ and gives $|x| \leq |y|$ if and only if $|x|^2 \leq |y|^2$. Finally $|x|^2 = x^2$, since $|x|$ is x or $-x$ and $(-x)^2 = x^2$, and likewise $|y|^2 = y^2$. This proves that $|x| \leq |y|$ if and only if $x^2 \leq y^2$.

2.5 Proof by Contradiction and the Classical Rules

Every rule so far either builds a statement or takes one apart. The rules for negation are no exception: ¬I (Definition 2.2.1) builds $\neg\varphi$ from a box that opens with φ and ends in $\bot$, and ¬E (Definition 2.1.1) uses $\neg\varphi$ against φ. Yet in three places this chapter has stopped one step short of a goal. This section adds the one rule that takes a negation away, reductio ad absurdum. It then derives from it the classical laws that mathematics uses without noticing, and shows that three of them are one principle in three forms.

2.5.1 Two boxes that look alike

Here are two arguments that a mathematician would both call "proof by contradiction".

The first is the proof of Theorem 2.2.1. It fixes a rational number r, supposes $r^2 = 2$, reaches a contradiction, and concludes that $r^2 \neq 2$. In the language of Definition 2.2.1, it opens a box with the assumption P ("$r^2 = 2$"), reaches $\bot$ inside it, and closes the box to get $\neg P$. That is ¬I.

The second is about a real number.

Example. A number whose square is at most itself

Let x be a real number with $x^2 \leq x$. Prove that $x \leq 1$.

Solution. Suppose, for a contradiction, that $x \leq 1$ fails, that is, $x > 1$. Then $x > 0$, and multiplying $x > 1$ by x gives $x^2 > x$. This contradicts $x^2 \leq x$. Hence $x \leq 1$.

Look at its shape. The box opens with the assumption $\neg P$, where P is "$x \leq 1$", reaches $\bot$, and closes to give P. If we tried to justify the last line by $\neg$I, we would get the negation of the box's assumption, $\neg\neg P$: "it is not the case that $x \leq 1$ fails". That is not P. The two arguments put their negations in different places. $\neg$I puts a $\neg$ **into** the conclusion; the second argument takes one **out**.

2.5.2 Where the rules so far stop short

The second argument is not a curiosity. Three times already in this chapter a derivation has stopped short in the same way: it reached a negation where the goal was not one.

1. Section 2.2 noted that a box opened with $\neg(x > 0)$ and closed by $\neg$I gives $\neg\neg(x > 0)$, and no rule so far gets from there to $x > 0$.
2. If you did Exercise C1 of Section 2.3, you met a second: Exercise (C1: Two premises that cannot both fail, used later) led from $P \to Q$ and $\neg P \to R$ to $\neg(\neg Q \wedge \neg R)$, and not to $Q \vee R$.
3. Section 2.4 noted, after Example (Negating both sides of an iff), that its converse, $\neg P \leftrightarrow \neg Q \vdash P \leftrightarrow Q$, resists the rules so far. The natural box on $\neg Q$, inside the box of P, closes to $\neg\neg Q$, not Q.

Try the smallest case, P from $\neg\neg P$. The goal P is a letter, so no introduction rule applies: the goal-directed moves have nothing to say. The premise $\neg\neg P$ is a negation, and a negation can be used only by $\neg$E, against the statement it negates, $\neg P$. So we assume $\neg P$ in a box. The box reaches $\bot$ at once, and $\neg$I closes it with $\neg\neg P$: the premise again. Every attempt goes round this circle.

That no derivation of P from $\neg\neg P$ exists with the rules so far is a claim about **all** derivations, which no search can settle; this section's Looking-ahead box says where it is proved. Here we name the missing step and add it as a rule.

2.5.3 Reductio ad absurdum

To prove φ, it is enough to show that "not φ" leads to a contradiction.

Definition 2.5.1. Reductio ad Absurdum

i	|	$\neg\varphi$	asm, for RAA
	|	$\vdots$	
j	|	$\bot$	
k	φ		RAA i–j

This is the rule RAA (**reductio ad absurdum**, "reduction to the absurd"). The line $\neg\varphi$ is an assumption, which opens the box and is **discharged** when the box closes. The line $\bot$ is a line **of that box itself**, not of a box inside it.

As in every schema (Definition 1.2.1), φ stands for any statement and i, j, k for line numbers. This rule has conditions that matter; read each one.

1. **The assumption is the exact negation of the goal.** To prove φ, the box opens with $\neg\varphi$: the negation of the **whole** of φ. If φ is $Q \vee R$, the box opens with $\neg(Q \vee R)$, not with

$\neg Q \vee R$. In a finished derivation this assumption is printed **"asm, for** RAA**"**; in prose it is the sentence "suppose, for a contradiction, that φ fails".

2. **The box must reach** $\bot$, as a line of the box itself: a $\bot$ inside a smaller box depends on a further assumption, as in $\neg$I.
3. **The result is** φ **itself**, not $\neg\neg\varphi$, which $\neg$I would give from the same box. This clause is the whole content of the rule.
4. **The box closes**, its assumption is discharged, and the justification cites the box as a range i–j.

In words: the box is a thought experiment in which φ fails. If the experiment ends in $\bot$, then φ cannot fail, and RAA lets us write φ. Set the new rule beside $\neg$I:

Rule	The box opens with	The box ends with	After the box
$\neg$I	φ	$\bot$	$\neg\varphi$
RAA	$\neg\varphi$	$\bot$	φ

The rows differ only in where the $\neg$ sits, and that makes RAA different in kind: every other primitive rule is the introduction or elimination rule of one connective ($\bot$E counts as the elimination rule of $\bot$, and R is the repetition clause of Definition 1.4.2), while RAA applies to every statement, whatever its main connective, even a letter.

In the Givens/Goal table, RAA adds $\neg\varphi$ to the Givens and makes $\bot$ the goal.

With the convention of Section 2.1, the argument of Example (A number whose square is at most itself) is this derivation.

1	$x^2 \le x$	
2	$\quad \neg(x \le 1)$	asm, for RAA
3	$\quad x > 1$	algebra 2
4	$\quad \neg(x^2 \le x)$	algebra 3
5	$\quad \bot$	$\neg$E 1, 4
6	$x \le 1$	RAA 2–5

Warning. The assumption that opens an RAA box must be the **exact** negation of the goal. To prove $x \le 1$ by contradiction, assume $x > 1$, since for real numbers "not $x \le 1$" says $x > 1$. Assuming $x \ge 1$ instead is a common slip. It is not the negation: $x = 1$ satisfies both $x \le 1$ and $x \ge 1$. A box opened with $x \ge 1$ cannot honestly reach $\bot$ while $x = 1$ is possible, so a "contradiction" that seems to be reached hides an error.

Quick check. A box opens with the assumption $\neg(x < 0)$ and reaches $\bot$, as a line of the box itself. Which of the following may be written after it closes, and by which rule? (i) $x < 0$; (ii) $\neg\neg(x < 0)$; (iii) $x \ge 0$.

Solution.

(i) By RAA: the box opened with the negation of $x < 0$. (ii) By ¬I: ¬I writes the negation of the box's assumption, and the assumption is $\neg(x < 0)$. (iii) By neither. For real numbers $x \geq 0$ says the same as $\neg(x < 0)$, which is exactly the statement the box has refuted.

Quick check. In the derivation of Example (A number whose square is at most itself) above, which line is the RAA step, which box does it cite, and what would ¬I have written there instead?

Solution. Line 6 is the RAA step, and it cites the box 2–5. ¬I would have written $\neg\neg(x \leq 1)$.

2.5.4 The move: when stuck, assume the negation of the goal

Chapter 1 and Sections 2.1–2.4 built two lists of moves. From the goal: its main connective chooses an introduction rule and so a box or two subgoals. From the premises: each available line's main connective chooses the elimination rule that uses it, and a goal $\bot$ asks for a clash. The new rule adds one move, for the moment when both lists are used up.

We are **stuck** at a gap when the goal-directed moves give nothing (the goal is a letter, or a disjunction with neither side in hand) and no premise-directed move brings the goal closer. Then open an RAA box on the negation of the goal. The new goal is $\bot$, so we hunt for a clash, and the box has given us one more line to clash with: $\neg\varphi$ itself.

When stuck, assume the negation of the goal, and hunt for a clash.

The move is a last resort, not a first one: try every other move first. Four of the derivations under the next heading reach the moment of being stuck, and their Ideas say where.

When RAA earns its place. Look ahead at where the classical step enters the derivations below: a goal that is a letter with nothing to build it (Proposition 2.5.1), and a disjunction neither of whose sides follows from what is available, each side's step breaking on its own (Proposition 2.5.3, De Morgan's law (d)). Premises that say little and a goal that decides something: that is where to expect RAA. Exercise B6 meets a third such disjunction.

2.5.5 Derivations with the new rule

The stall, resolved. The next result is the one on which we got stuck.

Proposition 2.5.1. Double Negation

For every statement P:

(a) $P \vdash \neg\neg P$, by a derivation that does not use RAA;

(b) $\neg\neg P \vdash P$, by a derivation that uses RAA.

Idea. In (a) the goal $\neg\neg P$ is the negation of $\neg P$, so the goal-directed move of Section 2.2 opens a $\neg$I box with $\neg P$, and the premise clashes with it at once.

In (b) the goal P is a letter, so no introduction rule applies. The only premise is a negation, which can only clash; the attempt above showed that clashing it inside a $\neg$I box leads back to the premise. We are stuck, so we assume the negation of the goal: an RAA box on $\neg P$. The premise clashes with it at once.

Proof.

(a) This is the derivation of Example (Two negations from one premise) (a), which uses the rules of Definition 2.2.1 alone; we repeat it to set it beside (b).

1	P	
2	$\quad\neg P$	asm, for $\neg$I
3	$\quad\bot$	$\neg$E 1, 2
4	$\neg\neg P$	$\neg$I 2–3

(b) An RAA box on $\neg P$:

1	$\neg\neg P$	
2	$\quad\neg P$	asm, for RAA
3	$\quad\bot$	$\neg$E 2, 1
4	P	RAA 2–3

Line 3 applies $\neg$E to the statement $\neg P$ on line 2 and to its negation $\neg\neg P$ on line 1. This proves both parts.

■

In both derivations, lines 2 and 3 form a box that opens with $\neg P$ and reaches $\bot$ by a clash with line 1. Only the rule that closes the box differs, and with it the last line. So the two directions are not mirror images: one needs only the rules for $\neg$, and the derivation of the other uses the new rule.

↻ **Second reading** · *Edge cases of reductio*

The degenerate case. The smallest use of RAA proves P from P.

Idea. There is nothing to find: the premise is the goal. The derivation is here only to show the rule at its smallest. The box opens with $\neg P$, and the one line it needs, $\bot$, comes from the clash of the assumption with the premise.

1	P	
2	$\quad \neg P$	asm, for RAA
3	$\quad \bot$	¬E 1, 2
4	P	RAA 2–3

It is legal and pointless, since line 1 already is the goal. It is the purest case of a pattern that turns up in real proofs: a contradiction box whose assumption does nothing except clash with a line we already had. A warning later in this section, after Example (A number below every positive number), returns to it.

Nonexample by minimal change. Change only the last line of (a), from $\neg\neg P$ to P, keeping the justification ¬I 2–3. Lines 1–3 still check. What fails is the conclusion of ¬I, which must be the negation of the box's assumption, here $\neg\neg P$. To write P, the justification must be RAA.

From now on both parts are derived rules in the sense of Definition 2.3.1. A line $\neg\neg\varphi$ may be justified by "¬¬I i", citing a line φ, and a line φ by "¬¬E i", citing a line $\neg\neg\varphi$. The first stall of the list above is now one line: from $\neg\neg(x > 0)$, ¬¬E gives $x > 0$.

Contraposition, backwards. Proposition 2.3.2 derives $\neg Q \to \neg P$ from $P \to Q$ with the rules for $\to$ and $\neg$ alone. In practice we use it the other way round. To prove "if n^2 is even then n is even", one may prove "if n is not even then n^2 is not even", and the step from that back to the original goes from $\neg Q \to \neg P$ to $P \to Q$.

Proposition 2.5.2. Converse of the Contrapositive

For all statements P and Q, $\neg Q \to \neg P \vdash P \to Q$.

Idea. The goal is an implication, so the goal-directed move opens a box with P, and the new goal is Q, a letter. The premise needs $\neg Q$ before it gives anything, and $\neg Q$ is nowhere; the box's assumption P is a letter and cannot be taken apart. We are stuck at Q, so we assume its negation, $\neg Q$. That is exactly what the premise wanted: $\to$E gives $\neg P$, which clashes with the P of the outer box, and RAA closes the inner box with Q.

In pictures: the stall at the letter Q; then the RAA box on $\neg Q$.

1	$\neg Q \to \neg P$	
2	$\quad P$	asm, for $\to$I
	$\quad \vdots$	
6	$\quad Q$	
7	$P \to Q$	$\to$I 2–6

1	$\neg Q \to \neg P$	
2	$\quad P$	asm, for $\to$I
3	$\quad \neg Q$	asm, for RAA
	$\quad \vdots$	
5	$\quad \bot$	
6	$\quad Q$	RAA 3–5
7	$P \to Q$	$\to$I 2–6

Proof. The derivation is:

1	$\neg Q \to \neg P$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad \neg Q$	asm, for RAA
4	$\quad\quad \neg P$	$\to$E 1, 3
5	$\quad\quad \bot$	$\neg$E 2, 4
6	$\quad Q$	RAA 3–5
7	$P \to Q$	$\to$I 2–6

Line 5 may cite line 2 because line 2 lies in a box that is still open. RAA enters at line 6. This proves the proposition.

■

So contraposition behaves like double negation: going from $P \to Q$ to $\neg Q \to \neg P$ needs no RAA, and our derivation of the way back uses one RAA.

Excluded middle. Every integer is even or odd. Every real number is ≥ 0 or < 0. The purely logical form of such facts is the law $P \vee \neg P$, which the next result derives from no premises at all.

Proposition 2.5.3. Excluded Middle

For every statement P, $\vdash P \vee \neg P$.

Idea. The goal is a disjunction, so Section 1.5's move postpones it: choose a side only when that side is in hand. But there are no premises, so no premise-directed move applies, and neither side is worth aiming at: P false breaks the step to P, and P true breaks the step to $\neg P$. We are stuck, so we assume the negation of the whole goal, $\neg(P \vee \neg P)$, which says that neither side holds, and hunt for a clash. The new line can clash only with $P \vee \neg P$, which $\vee$I builds from either side. So we manufacture a side, as in Example (Manufacturing the other half of a clash). If P held, then $P \vee \neg P$ would hold, against the assumption; so $\neg P$, by $\neg$I. But then $P \vee \neg P$ holds after all, and that is the final clash.

In pictures: the RAA box; then the final clash, written backwards.

1	$\quad \neg(P \vee \neg P)$	asm, for RAA
	$\quad \vdots$	
7	$\quad \bot$	
8	$P \vee \neg P$	RAA 1–7

1	$\neg(P \vee \neg P)$	asm, for RAA
	⋮	
5	$\neg P$	
6	$P \vee \neg P$	$\vee$I 5
7	$\bot$	$\neg$E 6, 1
8	$P \vee \neg P$	RAA 1–7

Proof. The derivation is:

1	$\neg(P \vee \neg P)$	asm, for RAA
2	P	asm, for $\neg$I
3	$P \vee \neg P$	$\vee$I 2
4	$\bot$	$\neg$E 3, 1
5	$\neg P$	$\neg$I 2–4
6	$P \vee \neg P$	$\vee$I 5
7	$\bot$	$\neg$E 6, 1
8	$P \vee \neg P$	RAA 1–7

The derivation has no premises, so it shows $\vdash P \vee \neg P$. RAA is used once, at line 8. This proves the proposition.

■

The assumption of line 1 is used twice, at lines 4 and 7. The first use refutes P; the second uses that refutation. Excluded middle now becomes a derived rule with no premises: a line $\varphi \vee \neg\varphi$ may be written anywhere, for any statement φ, justified by LEM with no line cited. Pasting in the eight lines above would replace it.

Its use is **proof by cases on a statement**. To derive χ, derive it once in a box that assumes φ and once in a box that assumes $\neg\varphi$, and combine the two boxes with the line $\varphi \vee \neg\varphi$ by $\vee$E. The fold after De Morgan's laws below uses it that way, and so does part (b) of the result after them.

A law with no negation. The last is a law in which $\neg$ does not occur at all.

Proposition 2.5.4. Peirce's Law

For all statements P and Q, $\vdash ((P \to Q) \to P) \to P$.

In words: a law with no $\neg$ in it, whose derivation here still uses RAA. *The derivation is in the fold below.*

Second reading · *Peirce's law, derived*

Try it first (ten minutes): use the goal-directed and premise-directed moves of the earlier sections on this goal until they stall. At which gap do they stall, and which line, if it were available there, would let you finish?

Idea. The goal is an implication, so we open a box on $(P \to Q) \to P$ with the goal P, a letter. The only line that can produce P is the assumption itself, by $\to$E, once we have its hypothesis $P \to Q$. So $P \to Q$ becomes a subgoal, and the goal-directed move opens a box on P with the goal Q. There the moves stall: the resources are $(P \to Q) \to P$ and P, and neither says anything about Q. What the stalled gap lacks is a clash, which $\bot$E would turn into Q.

We are stuck, so the new move applies, but where? Assuming $\neg Q$ at the stalled gap gives nothing new: a clash with $\neg Q$ needs Q, the stalled goal itself. So apply the move one level out, to the goal P, and assume $\neg P$. That is exactly what the stalled gap lacked: inside the box of P, the pair P, $\neg P$ gives $\bot$, and $\bot$E gives Q. Then $P \to Q$, then P by $\to$E, which clashes with $\neg P$.

Proof. The derivation is:

1	$(P \to Q) \to P$	asm, for $\to$I
2	$\quad \neg P$	asm, for RAA
3	$\quad\quad P$	asm, for $\to$I
4	$\quad\quad \bot$	$\neg$E 3, 2
5	$\quad\quad Q$	$\bot$E 4
6	$\quad P \to Q$	$\to$I 3–5
7	$\quad P$	$\to$E 1, 6
8	$\quad \bot$	$\neg$E 7, 2
9	P	RAA 2–8
10	$((P \to Q) \to P) \to P$	$\to$I 1–9

The derivation has no premises, and RAA enters at line 9. This proves the proposition. ■

The assumption $\neg P$ of line 2 is used twice: at line 4 it makes the innermost box contradictory, and at line 8 it clashes with the P that line 7 obtained. Notice where the new box went: around the stalled box, not inside it. An assumption made one level out is available in every gap inside its box, and so it could give the stalled gap the line it lacked. Peirce's law contains no $\neg$, yet our derivation uses RAA: the classical rule is about more than negation.

2.5.6 De Morgan's laws

The negation of an "and" is an "or" of negations, and the negation of an "or" is an "and" of negations. For example, a real number x fails to lie in the interval $[0, 1]$, that is, $\neg(0 \leq x \wedge x \leq 1)$, exactly when $x < 0$ or $x > 1$. Each law has two directions, so there are four derivations to give.

Proposition 2.5.5. De Morgan's Laws

For all statements P and Q:

(a) $\neg(P \vee Q) \vdash \neg P \wedge \neg Q$;
(b) $\neg P \wedge \neg Q \vdash \neg(P \vee Q)$;
(c) $\neg P \vee \neg Q \vdash \neg(P \wedge Q)$;
(d) $\neg(P \wedge Q) \vdash \neg P \vee \neg Q$.

Try it first (ten minutes): look for a derivation of (d) with the rules of the earlier sections. Which rule would give its last line, and what would that rule need?

Idea. Parts (a)–(c) yield to their goals. In (a) the goal is a conjunction of negations: two ¬I boxes, one per conjunct. In (b) and (c) the goal is a negation: one ¬I box, with the disjunction used inside it by ∨E.

In pictures, for (a): two subgoals; then a box under each.

1	$\neg(P \vee Q)$	
	⋮	
5	$\neg P$	
	⋮	
9	$\neg Q$	
10	$\neg P \wedge \neg Q$	$\wedge$I 5, 9

1	$\neg(P \vee Q)$	
2	P	asm, for ¬I
	⋮	
4	$\bot$	
5	$\neg P$	¬I 2–4
6	Q	asm, for ¬I
	⋮	
8	$\bot$	
9	$\neg Q$	¬I 6–8
10	$\neg P \wedge \neg Q$	$\wedge$I 5, 9

For (b): the box on $P \vee Q$; then the two cases inside it.

1	$\neg P \wedge \neg Q$	
2	$\neg P$	$\wedge$E 1
3	$\neg Q$	$\wedge$E 1
4	$P \vee Q$	asm, for $\neg$I
	⋮	
9	$\bot$	
10	$\neg(P \vee Q)$	$\neg$I 4–9

1	$\neg P \wedge \neg Q$	
2	$\neg P$	$\wedge$E 1
3	$\neg Q$	$\wedge$E 1
4	$P \vee Q$	asm, for $\neg$I
5	P	asm, for $\vee$E
	⋮	
6	$\bot$	
7	Q	asm, for $\vee$E
	⋮	
8	$\bot$	
9	$\bot$	$\vee$E 4, 5–6, 7–8
10	$\neg(P \vee Q)$	$\neg$I 4–9

Part (d) resists. Its goal is a disjunction, so the natural last line is $\vee$I, from $\neg P$ or from $\neg Q$. But the step from the premise to $\neg P$ can be broken: with P true and Q false the premise is true and $\neg P$ false; and the step to $\neg Q$ breaks with P false and Q true. So neither side is worth aiming at. The premise itself can be used only by $\neg$E, against $P \wedge Q$, which we do not have. We are stuck at the whole goal, so we assume its negation, $\neg(\neg P \vee \neg Q)$, which says that neither $\neg P$ nor $\neg Q$ holds. It should give P and Q. Each is a letter, so each gets a small RAA box of its own, whose clash is manufactured by $\vee$I. Together P and Q clash with the premise.

For (d): the RAA box on the negated goal; then a box for each of P and Q.

1	$\neg(P \wedge Q)$	
2	$\neg(\neg P \vee \neg Q)$	asm, for RAA
	⋮	
11	$P \wedge Q$	
12	$\bot$	$\neg$E 11, 1
13	$\neg P \vee \neg Q$	RAA 2–12

1	$\neg(P \wedge Q)$	
2	$\quad \neg(\neg P \vee \neg Q)$	asm, for RAA
3	$\quad\quad \neg P$	asm, for RAA
	$\quad\quad \vdots$	
5	$\quad\quad \bot$	
6	$\quad P$	
7	$\quad\quad \neg Q$	asm, for RAA
	$\quad\quad \vdots$	
9	$\quad\quad \bot$	
10	$\quad Q$	
11	$\quad P \wedge Q$	$\wedge$I 6, 10
12	$\quad \bot$	$\neg$E 11, 1
13	$\neg P \vee \neg Q$	RAA 2–12

Proof.

(a) Two $\neg$I boxes, one for each conjunct, then $\wedge$I:

1	$\neg(P \vee Q)$	
2	$\quad P$	asm, for $\neg$I
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad \bot$	$\neg$E 3, 1
5	$\neg P$	$\neg$I 2–4
6	$\quad Q$	asm, for $\neg$I
7	$\quad P \vee Q$	$\vee$I 6
8	$\quad \bot$	$\neg$E 7, 1
9	$\neg Q$	$\neg$I 6–8
10	$\neg P \wedge \neg Q$	$\wedge$I 5, 9

(b) One $\neg$I box that opens with $P \vee Q$. Inside it, $\vee$E splits into two cases, and each case clashes with one half of the premise:

1	$\neg P \wedge \neg Q$	
2	$\neg P$	$\wedge$E 1
3	$\neg Q$	$\wedge$E 1
4	$P \vee Q$	asm, for $\neg$I
5	P	asm, for $\vee$E
6	$\bot$	$\neg$E 5, 2
7	Q	asm, for $\vee$E
8	$\bot$	$\neg$E 7, 3
9	$\bot$	$\vee$E 4, 5–6, 7–8
10	$\neg(P \vee Q)$	$\neg$I 4–9

(c) The same shape as (b), with the roles of $\wedge$ and $\vee$ exchanged:

1	$\neg P \vee \neg Q$	
2	$P \wedge Q$	asm, for $\neg$I
3	P	$\wedge$E 2
4	Q	$\wedge$E 2
5	$\neg P$	asm, for $\vee$E
6	$\bot$	$\neg$E 3, 5
7	$\neg Q$	asm, for $\vee$E
8	$\bot$	$\neg$E 4, 7
9	$\bot$	$\vee$E 1, 5–6, 7–8
10	$\neg(P \wedge Q)$	$\neg$I 2–9

(d) RAA on the goal, and inside it two smaller RAA boxes:

1	$\neg(P \wedge Q)$	
2	$\quad \neg(\neg P \vee \neg Q)$	asm, for RAA
3	$\quad\quad \neg P$	asm, for RAA
4	$\quad\quad \neg P \vee \neg Q$	$\vee$I 3
5	$\quad\quad \bot$	$\neg$E 4, 2
6	$\quad P$	RAA 3–5
7	$\quad\quad \neg Q$	asm, for RAA
8	$\quad\quad \neg P \vee \neg Q$	$\vee$I 7
9	$\quad\quad \bot$	$\neg$E 8, 2
10	$\quad Q$	RAA 7–9
11	$\quad P \wedge Q$	$\wedge$I 6, 10
12	$\quad \bot$	$\neg$E 11, 1
13	$\neg P \vee \neg Q$	RAA 2–12

The derivations of (a), (b), and (c) close every box by $\neg$I or $\vee$E and use no RAA. The derivation of (d) uses RAA at lines 6, 10, and 13. This proves all four laws.

■

Laws (a)–(c) conclude negations or a conjunction of negations, and $\neg$I builds negations. Law (d) concludes a **disjunction**, and $\vee$I needs to know **which** side holds, while the premise says only that P and Q do not both hold. RAA gets round the missing decision without making it.

Second reading · *De Morgan (d) by excluded middle*

Excluded middle makes the decision instead. Split on $P \vee \neg P$. In the case $\neg P$, $\vee$I gives $\neg P \vee \neg Q$ at once. In the case P, a $\neg$I box on Q builds $P \wedge Q$, which clashes with the premise, giving $\neg Q$; then $\vee$I gives $\neg P \vee \neg Q$. Written out, this route is 11 lines, with LEM as its one classical line. Both routes use a classical rule.

Quick check. Suppose RAA were deleted from the list of rules. Which of the four derivations in the proof of Proposition 2.5.5 would still be correct, and what is the first line of the others that would lose its justification?

Solution. The derivations of (a), (b), and (c) would still be correct: they use $\neg$I, $\neg$E, $\vee$I, $\vee$E, $\wedge$I, and $\wedge$E only. In the derivation of (d), line 6 is the first to lose its justification: it closes the box that opened with $\neg P$ and writes P, which is RAA. (Lines 10 and 13 lose theirs too.)

2.5.7 Implication as a disjunction

In everyday mathematics "if P then Q" and "either P fails or Q holds" are used interchangeably. One direction of that exchange needs only the earlier rules; our derivation of the other uses a classical rule.

Proposition 2.5.6. Implication as a Disjunction

For all statements P and Q:

(a) $\neg P \vee Q \vdash P \to Q$, by a derivation that does not use RAA;

(b) $P \to Q \vdash \neg P \vee Q$, by a derivation that uses LEM.

Idea.

(a) The premise is a disjunction, so we use it by cases, each ending in the current goal $P \to Q$. In each case that goal is an implication, so we open a box with P. In the case $\neg P$, the box clashes at once and $\bot$E gives Q: this is the derivation of Proposition 2.1.2. In the case Q, the box repeats Q.
In pictures, for (a): the two cases; then a box on P in each.

1	$\neg P \vee Q$	
2	$\quad \neg P$	asm, for $\vee$E
	$\quad \vdots$	
6	$\quad P \to Q$	
7	$\quad Q$	asm, for $\vee$E
	$\quad \vdots$	
10	$\quad P \to Q$	
11	$P \to Q$	$\vee$E 1, 2–6, 7–10

1	$\neg P \vee Q$	
2	$\quad \neg P$	asm, for $\vee$E
3	$\quad\quad P$	asm, for $\to$I
	$\quad\quad \vdots$	
5	$\quad\quad Q$	
6	$\quad P \to Q$	$\to$I 3–5
7	$\quad Q$	asm, for $\vee$E
8	$\quad\quad P$	asm, for $\to$I
	$\quad\quad \vdots$	
9	$\quad\quad Q$	
10	$\quad P \to Q$	$\to$I 8–9
11	$P \to Q$	$\vee$E 1, 2–6, 7–10

(b) The goal is a disjunction, so we postpone it. The premise is an implication, which gives Q only once we have P, and we do not. So we are stuck, and this time excluded middle is the better tool than an RAA box: we decide P first. In the case P the premise gives Q; in the case $\neg P$ the left side is in hand.

For (b): the split; then $\to$E in the case P.

1	$P \to Q$	
2	$P \vee \neg P$	LEM
3	$\quad P$	asm, for $\vee$E
	$\quad \vdots$	
5	$\quad \neg P \vee Q$	
6	$\quad \neg P$	asm, for $\vee$E
	$\quad \vdots$	
7	$\quad \neg P \vee Q$	
8	$\neg P \vee Q$	$\vee$E 2, 3–5, 6–7

1	$P \to Q$	
2	$P \vee \neg P$	LEM
3	$\quad P$	asm, for $\vee$E
4	$\quad Q$	$\to$E 1, 3
	$\quad \vdots$	
5	$\quad \neg P \vee Q$	
6	$\quad \neg P$	asm, for $\vee$E
	$\quad \vdots$	
7	$\quad \neg P \vee Q$	
8	$\neg P \vee Q$	$\vee$E 2, 3–5, 6–7

Proof.

(a) $\vee$E on the premise, with an implication built in each case:

1	$\neg P \lor Q$	
2	$\quad \neg P$	asm, for $\lor$E
3	$\quad\quad P$	asm, for $\to$I
4	$\quad\quad \bot$	$\neg$E 3, 2
5	$\quad\quad Q$	$\bot$E 4
6	$\quad P \to Q$	$\to$I 3–5
7	$\quad Q$	asm, for $\lor$E
8	$\quad\quad P$	asm, for $\to$I
9	$\quad\quad Q$	R 7
10	$\quad P \to Q$	$\to$I 8–9
11	$P \to Q$	$\lor$E 1, 2–6, 7–10

Line 9 repeats line 7 inside the box of line 8, because $\to$I needs Q as a line of that box itself.

(b) Cases on P:

1	$P \to Q$	
2	$P \lor \neg P$	LEM
3	$\quad P$	asm, for $\lor$E
4	$\quad Q$	$\to$E 1, 3
5	$\quad \neg P \lor Q$	$\lor$I 4
6	$\quad \neg P$	asm, for $\lor$E
7	$\quad \neg P \lor Q$	$\lor$I 6
8	$\neg P \lor Q$	$\lor$E 2, 3–5, 6–7

Line 2 is Proposition 2.5.3 used as a derived rule, and the only RAA is inside it. This proves both parts.

■

The two routes of the section now stand side by side. RAA assumes the goal fails and hunts for a clash; LEM splits on a statement of our choosing and gives each case a resource it lacked. When a premise is waiting for a statement S (here $P \to Q$ waits for P), splitting on S is usually the shorter route.

2.5.8 One principle in three forms

RAA gave us $\neg\neg$E and excluded middle. Had we added one of those two as the new primitive rule instead of RAA, would we have proved less?

To say this exactly, we need two rule forms and one notion. $\neg\neg$E **taken as a rule** says: if $\neg\neg\varphi$ stands on an available line, we may write φ. LEM **taken as a rule** says: we may write $\varphi \lor \neg\varphi$

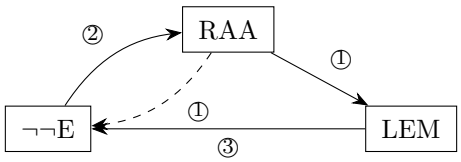
on any line, for any statement φ, citing nothing. Given a list of rules, a rule X **derives** a rule Y over that list if every use of Y in a derivation can be replaced by a fragment that uses only X and the rules of the list, without changing the premises or the last line. Then every sequent derivable with Y added to the list is derivable with X added instead.

Theorem 2.5.7. The Classical Rules Are Interderivable

Over the primitive rules other than RAA, each of the three rules RAA, ¬¬E, and LEM derives the other two. In particular, adding any one of the three to the other primitive rules gives the same derivable sequents.

In words: RAA, ¬¬E *and* LEM *are one principle in three forms; adding any one of them to the other rules proves exactly what adding any other does.*

In the picture, an arrow from X to Y means that X derives Y over the other primitive rules; going round the solid cycle, each rule derives the other two. ① is Proposition 2.5.3 on the solid arrow and Proposition 2.5.1 (b) on the dashed one; ② closes the box by ¬I, then applies ¬¬E; ③ splits on $\varphi \vee \neg\varphi$.



Part ② says that a step of RAA can be replaced by ¬I followed by ¬¬E, keeping the box. On the derivation of Example (A number whose square is at most itself), line 6 becomes two lines:

1	$x^2 \le x$	
2	$\quad \neg(x \le 1)$	asm, for ¬I
3	$\quad x > 1$	algebra 2
4	$\quad \neg(x^2 \le x)$	algebra 3
5	$\quad \bot$	¬E 1, 4
6	$\neg\neg(x \le 1)$	¬I 2–5
7	$x \le 1$	¬¬E 6

Line 6 is correct by ¬I: the box opens with $\neg(x \le 1)$, and $\neg\neg(x \le 1)$ is the negation of that assumption. The box itself is unchanged: both rules ask for $\bot$ as a line of the box, and differ only in what they write after it. Part ③, shown on the same example in the fold below, replaces that ¬¬E line by a split written by LEM.

So the classical part of our rules is **one** extra principle, whichever of its three forms we write down.

↻ **Second reading** · *Part ③ on the example, and the proof: replacing one rule by another*

Part ③ on the example. Part ③ replaces the ¬¬E line of ② by a split on $(x \le 1) \vee \neg(x \le 1)$, written by LEM:

1	$x^2 \leq x$	
2	$\quad \neg(x \leq 1)$	asm, for ¬I
3	$\quad x > 1$	algebra 2
4	$\quad \neg(x^2 \leq x)$	algebra 3
5	$\quad \bot$	¬E 1, 4
6	$\neg\neg(x \leq 1)$	¬I 2–5
7	$(x \leq 1) \vee \neg(x \leq 1)$	LEM
8	$\quad x \leq 1$	asm, for ∨E
9	$\quad \neg(x \leq 1)$	asm, for ∨E
10	$\quad \bot$	¬E 9, 6
11	$\quad x \leq 1$	⊥E 10
12	$x \leq 1$	∨E 7, 8–8, 9–11

The first case is the goal; the second clashes with line 6 and closes itself. Together, ② and ③ turned the one RAA line into steps of ¬I, LEM and the rules of the earlier sections: that is how LEM derives RAA.

Idea. Replacements can be chained, so it is enough to go once round a cycle. ① RAA derives the other two: that is Proposition 2.5.3 and Proposition 2.5.1 (b). ② ¬¬E derives RAA: ¬I closes an RAA box, which runs from $\neg\varphi$ to $\bot$, with $\neg\neg\varphi$, and ¬¬E finishes. ③ LEM derives ¬¬E: split on $\varphi \vee \neg\varphi$; the case φ is the goal, and the case $\neg\varphi$ clashes with $\neg\neg\varphi$.

Before reading on: in the fragment for ③ below, which rule does line $k+4$ use, and why may line $k+3$ cite line m?

Proof. ① Suppose a derivation uses ¬¬E at line k, writing φ and citing a line m that holds $\neg\neg\varphi$. Replace line k by the last three lines of the derivation in Proposition 2.5.1 (b), with its premise replaced by line m: a box opening with $\neg\varphi$, the line $\bot$ by ¬E citing the assumption and line m, and then φ by RAA. A use of LEM, writing $\varphi \vee \neg\varphi$, is replaced by the eight lines of the derivation in Proposition 2.5.3, with φ for P; that derivation has no premises. So RAA derives ¬¬E and LEM.

② Suppose a derivation uses RAA at line k, closing a box, lines i to j, that opens with $\neg\varphi$ and ends with $\bot$. Keep the box and replace line k by two lines:

i	$\quad \neg\varphi$	asm, for ¬I
	$\quad \vdots$	
j	$\quad \bot$	
k	$\neg\neg\varphi$	¬I i–j
$k+1$	φ	¬¬E k

Line k is correct by ¬I, since the box opens with $\neg\varphi$ and $\neg\neg\varphi$ is the negation of that

assumption. Both rules ask for $\bot$ as a line of the box itself, so the box needs no change. So $\neg\neg$E derives RAA.

③ Suppose a derivation uses $\neg\neg$E at line k, writing φ and citing a line m that holds $\neg\neg\varphi$. Replace line k by six lines:

m	$\neg\neg\varphi$	
	⋮	
k	$\varphi \vee \neg\varphi$	LEM
$k+1$	$\quad\vert\ \varphi$	asm, for $\vee$E
$k+2$	$\quad\vert\ \neg\varphi$	asm, for $\vee$E
$k+3$	$\quad\vert\ \bot$	$\neg$E $k+2$, m
$k+4$	$\quad\vert\ \varphi$	$\bot$E $k+3$
$k+5$	φ	$\vee$E k, $k+1$–$k+1$, $k+2$–$k+4$

The first case is a one-line box: it opens with φ and so already ends with the goal. In the second case, line $k+3$ may cite line m, because line m was available at the old line k and every new line sits in the same boxes as the old line k, or in boxes opened after it. No rule other than LEM and the rules of the list is used. So LEM derives $\neg\neg$E.

In each replacement the later lines are renumbered, and a later citation of the old line k now cites the fragment's last line, which holds the same statement in the same boxes. The premises and the last line of the derivation are unchanged.

Finally, replacements can be chained. LEM derives RAA: replace each use of RAA by the fragment of ②, and then each use of $\neg\neg$E so created by the fragment of ③. $\neg\neg$E derives LEM: replace each use of LEM by the derivation of ①, and then each use of RAA in it by the fragment of ②. With ①, ②, and ③, this shows that each of the three rules derives the other two, as claimed.

■

The base list matters: part ③ uses $\bot$E at line $k+4$, to close the impossible case as in Example (An impossible case).

↻ **Second reading** · *Other books, and a rule we could drop*

Remark. Some books call exactly our RAA "proof by contradiction" or PBC, and some take $\neg\neg$E as their primitive rule and derive RAA from it. By Theorem 2.5.7 nothing changes: all these choices prove the same sequents.

Remark. $\bot$E could have been dropped from the primitive rules, since RAA and repetition derive it (Exercise B3 below asks for the derivation). We keep it so that the primitive rules

other than RAA form a system of their own.

2.5.9 How reductio looks in a mathematical proof

Here is an argument of a kind that analysis uses constantly; read it one sentence at a time, with the rule of each step in brackets.

Example. A number below every positive number

Let x be a real number with $x \geq 0$, and suppose that $x \leq \varepsilon$ for **every** real number $\varepsilon > 0$. Prove that $x = 0$.

Solution. Suppose, for a contradiction, that $x \neq 0$. Since $x \geq 0$ and $x \neq 0$, we have $x > 0$. Then $\varepsilon := x/2$ is a real number with $\varepsilon > 0$, so the hypothesis applies to it and gives $x \leq x/2$. Subtracting $x/2$ from both sides gives $x/2 \leq 0$, and multiplying by 2 gives $x \leq 0$. This contradicts $x > 0$. Hence $x = 0$.

1. "Suppose, for a contradiction, that $x \neq 0$." [An RAA box opens, with $\neg(x = 0)$, the exact negation of the goal $x = 0$.]
2. "Since $x \geq 0$ and $x \neq 0$, we have $x > 0$." Example (An impossible case)
3. "Then $\varepsilon := x/2$ is a real number with $\varepsilon > 0$." [School algebra.]
4. "So the hypothesis applies to it and gives $x \leq x/2$." [The hypothesis is about **every** $\varepsilon > 0$; applying it to one is a step about "every", used as ordinary reasoning (the note at the start of this chapter).]
5. "... gives $x/2 \leq 0$, and ... $x \leq 0$." [School algebra.]
6. "This contradicts $x > 0$." [¬E: for real numbers $x > 0$ is $\neg(x \leq 0)$, the negation of the line $x \leq 0$.]
7. "Hence $x = 0$." [RAA closes the box.]

¬I or RAA? The phrase "suppose not" covers two rules. Here is one statement proved both ways, for a real number x: *if* $x^3 < 0$*, then* $x < 0$.

- **By the contrapositive (CP).** We prove that if $x \geq 0$, then $x^3 \geq 0$. Suppose $x \geq 0$ [assumption for →I]. Then $x^3 = x \cdot x \cdot x \geq 0$ [algebra]. So if $x \geq 0$, then $x^3 \geq 0$ [→I], and by contraposition, if x^3 is not ≥ 0, then x is not ≥ 0 [CP]: that is, if $x^3 < 0$, then $x < 0$.
- **By contradiction (RAA).** Suppose $x^3 < 0$ [assumption for →I]. Suppose, for a contradiction, that $x < 0$ fails, that is, $x \geq 0$ [an RAA box on $\neg(x < 0)$]. Then $x^3 \geq 0$, which contradicts $x^3 < 0$ [¬E]. Hence $x < 0$ [RAA], and so if $x^3 < 0$, then $x < 0$ [→I].

The first proof introduces negations and removes none; the second removes one, at its RAA step. Which rule a "suppose not" uses is decided by the lines as written: what the box assumes, and what is written after it. Under the convention of Section 2.1 an order statement can be written either way: with the goal $x < 0$ written as $\neg(x \geq 0)$, the box on $x \geq 0$ closes by ¬I; and Example (A number whose square is at most itself), with its goal written $\neg(x > 1)$ and its box on $x > 1$, would close by ¬I too.

The proof of Theorem 2.2.1 assumes the positive statement $r^2 = 2$ and concludes its negation, so it is ¬I.

Quick check. ¬I or RAA? (i) "Suppose 15 is prime. Then ... This is a contradiction. Hence 15 is not prime." (ii) "Suppose n is not odd. Then ... This is a contradiction. Hence n is odd."

Solution.

(i) ¬I: the box opens with the statement whose negation is concluded. (ii) RAA: the box opens with ¬(n is odd), and the conclusion is "n is odd".

Second reading · *RAA is never needed to finish a negation*

For one kind of goal the new rule is never needed at the last step. Suppose the goal is itself a negation $\neg\psi$. An RAA box for it opens with $\neg\neg\psi$. Replace it by a ¬I box that opens with ψ; inside it, derive $\neg\neg\psi$ from ψ by the three lines of Example (Two negations from one premise) (a) (a box on $\neg\psi$ that clashes with ψ); then continue as the RAA box did, citing that line wherever the RAA box cited its assumption. The box still ends in $\bot$, and ¬I closes it with $\neg\psi$. So the **last** step of a derivation of a negation never needs RAA, though steps inside its box still may. RAA earns its place on the other goals.

When Q is itself a negation, the inner goal in Proposition 2.5.2 is a negation, and its RAA can be replaced by ¬I in this way.

Warning. An RAA box on $\neg\varphi$ whose assumption is used **only at its last step**, against a φ derived without it, is legal, but it is a direct proof in disguise: the lines before the clash already derive φ, so the box can be deleted. The smallest case opens a box on $\neg P$ only to clash with a premise P that is already the goal. It is a direct proof in costume. (A box that reaches $\bot$ without using its assumption at all is a different matter: the same lines, written before the box, derive $\bot$ without it, and $\bot$E then gives the goal at once.)

Warning. Not every "proof by contradiction" is RAA. Calling the proof of Theorem 2.2.1 an RAA is a mislabeling: its box opens with $r^2 = 2$, not with a negation, and it concludes $r^2 \neq 2$, which ¬I writes. The difference matters when you count classical steps: the box of that proof closes by ¬I and contributes none.

2.5.10 The rule card

Before the card, the new rule on large statements; each box runs from the assumption shown to $\bot$.

Example. Instances of RAA

φ	The box opens with	New line
$P \vee \neg Q$	$\neg(P \vee \neg Q)$	$P \vee \neg Q$
$(P \to Q) \to R$	$\neg((P \to Q) \to R)$	$(P \to Q) \to R$
$\neg Q$	$\neg\neg Q$	$\neg Q$

The third row is legal; a ¬I box opened by Q would also do (the fold "RAA is never needed to finish a negation").

Not an instance. A box opens with $\neg P \vee \neg Q$ and reaches $\bot$. The line "$P \wedge Q$ — RAA" is not an instance: the assumption must be $\neg(P \wedge Q)$, the negation of the whole line written. $\neg P \vee \neg Q$ says the same only after a derivation (De Morgan's laws), and the rule compares symbols.

Chapter 1's seven rules are on the card at the end of Section 1.6; here are this chapter's so far.

Rule	Shape	You write	Lean
¬E	$\varphi, \neg\varphi \Longrightarrow \bot$	"…, which contradicts …"	`exact hn hp` (goal `False`)
⊥E	$\bot \Longrightarrow$ any φ	"This case cannot occur." "So there is nothing to prove."	`exfalso`
¬I	box φ … $\bot \Longrightarrow \neg\varphi$	"Suppose P. … This is a contradiction. Hence not P."	`intro hp` (goal `False`)
MT (derived)	$\varphi \to \psi, \neg\psi \Longrightarrow \neg\varphi$	"Since Q fails, so does P."	`mt h hnq`
CP (derived)	$\varphi \to \psi \Longrightarrow \neg\psi \to \neg\varphi$	"We prove the contrapositive: suppose not Q; … so not P."	`mt`
DS (derived)	$\varphi \vee \psi, \neg\varphi \Longrightarrow \psi$	"P is impossible, so Q."	`h.resolve_left hnp`
↔I	box φ … ψ, box ψ … $\varphi \Longrightarrow \varphi \leftrightarrow \psi$	"(⇒) Suppose P. … so Q. (⇐) Suppose Q. … so P."	`constructor`
↔E	$\varphi \leftrightarrow \psi, \varphi \Longrightarrow \psi$ (or with ψ, φ)	"Since P holds if and only if Q does, and P holds, so does Q."	`h.mp hp` / `h.mpr hq`
RAA	box $\neg\varphi$ … $\bot \Longrightarrow \varphi$	"Suppose, for a contradiction, that P fails. … This is impossible. Hence P."	`by_contra h`
¬¬E (derived)	$\neg\neg\varphi \Longrightarrow \varphi$	"It cannot be that P fails; so P."	`not_not.mp h`
LEM (derived)	nothing $\Longrightarrow \varphi \vee \neg\varphi$	"Either P or not. If P, … If not, … In either case, …"	`by_cases h : P`

The Lean column is for readers who use that proof assistant; nothing later depends on it.

Warning (Not allowed: RAA). Each step is wrong; the tag names the clause it breaks.

- Open the box for $Q \lor R$ with $\neg Q \lor R$: clause 1, the box opens with the negation of the **whole** goal.
- Write φ after a box opened by φ: clause 1 (the pair below).
- Close the box on a $\bot$ of a box inside it: clause 2.
- Write $\neg\neg\varphi$ by RAA: clause 3; $\neg\neg\varphi$ is the line ¬I writes.
- To prove $x \leq 1$, assume $x \geq 1$: clause 1, the negation of $x \leq 1$ is $x > 1$.

Clause 1 is where the rule bites. Premises $(P \lor R) \to Q$ and $\neg Q$; wrong on the left, right on the right:

1	$(P \lor R) \to Q$	
2	$\neg Q$	
3	$\quad P$	asm, for RAA
4	$\quad P \lor R$	∨I 3
5	$\quad Q$	→E 1, 4
6	$\quad \bot$	¬E 5, 2
7	P	RAA 3–6

1	$(P \lor R) \to Q$	
2	$\neg Q$	
3	$\quad P$	asm, for ¬I
4	$\quad P \lor R$	∨I 3
5	$\quad Q$	→E 1, 4
6	$\quad \bot$	¬E 5, 2
7	$\neg P$	¬I 3–6

The left box opened with P, not $\neg P$, so it refuted P and cannot conclude it. The claim $(P \lor R) \to Q,\ \neg Q \vdash P$ breaks: let x be a real number, with P: "$x > 2$", R: "$x < -2$", Q: "$x^2 > 4$". The first premise holds for every x, and at $x = 0$ the second holds and P fails.

Quick check. A box opens with $\neg(x \neq 2)$, for a real number x, and reaches $\bot$ as a line of the box itself. What may RAA write after it, and what may ¬I?

Solution. RAA writes $x \neq 2$; ¬I writes $\neg\neg(x \neq 2)$, that is $\neg\neg\neg(x = 2)$.

Remark (Looking ahead). Chapter 4 proves that P cannot be derived from $\neg\neg P$, nor Peirce's law from no premises, without RAA or one of its two equivalents, and states the same, without proof, for De Morgan's law (d) and for $(P \to Q) \lor (Q \to P)$. Chapter 8 tells the story of the mathematicians who declined the classical principle, and shows how to spot a proof by contradiction that is a direct proof in costume.

2.5.11 Exercises

2.5.11.1 A. Check your understanding

Exercise. A1

(a) "Suppose 7 is even. Then $7 = 2k$ for some integer k, so $k = 7/2$, which is not an integer. This is impossible. Hence 7 is not even." Is the last step ¬I, RAA, or neither? What did the box open with?

(b) "Let x be real. Suppose $x^2 \geq 2x - 1$ fails, that is, $x^2 < 2x - 1$. Then $(x-1)^2 < 0$, which is impossible. Hence $x^2 \geq 2x - 1$." Same questions.

(c) "Assume P. Then, as shown above, Q. Hence $P \to Q$." Same questions.

(d) State RAA and ¬I side by side, and say which of the two the derivation of Proposition 2.5.1 (a) uses.

(e) In the derivation of Proposition 2.5.5 (d), why is a last line ∨I, from $\neg P$ or $\neg Q$ derived from the premise alone, not worth aiming at?

(f) In the same derivation, the assumption $\neg(\neg P \vee \neg Q)$ on line 2 is cited twice. Name the two lines that cite it, and the rule each uses.

(g) In the instance of ② on the opening example (in "One principle in three forms"), which rule justifies the line $\neg\neg(x \leq 1)$, and why does the box of lines 2–5 need no change?

Solution.

(a) ¬I: the box opened with "7 is even", and the conclusion is its negation. ($7/2$ is not an integer because $3 < 7/2 < 4$.)

(b) RAA: the box opened with the negation of $x^2 \geq 2x-1$, and the conclusion is $x^2 \geq 2x-1$ itself. (The clash: $x^2 < 2x-1$ rearranges to $(x-1)^2 < 0$, and squares are nonnegative.)

(c) Neither: this is →I. The box opened with P and ended with Q; no $\bot$ appears.

(d) ¬I: a box that opens with φ and reaches $\bot$ gives $\neg\varphi$. RAA: a box that opens with $\neg\varphi$ and reaches $\bot$ gives φ. The derivation of Proposition 2.5.1 (a) uses ¬I, with $\neg P$ as the assumption, so its conclusion is the negation of $\neg P$, namely $\neg\neg P$.

(e) ∨I would need $\neg P$ or $\neg Q$ on a line derived from the premise alone, and each of those steps can be broken: with P true and Q false the premise is true and $\neg P$ false, and with P false and Q true the premise is true and $\neg Q$ false. So neither step is worth aiming at: each breaks (Section 1.1's test).

(f) Line 5 cites it, by ¬E against line 4, $\neg P \vee \neg Q$; and line 9 cites it, by ¬E against line 8, the same statement. Both lines lie in boxes inside the box of line 2, which is still open, so the citations are legal.

(g) ¬I: the box opens with $\neg(x \leq 1)$ and ends with $\bot$, so ¬I gives the negation of its assumption, $\neg\neg(x \leq 1)$. The box needs no change because ¬I, like RAA, asks only for a box that ends with $\bot$ as a line of the box itself; the two rules differ in what they write after the box, not in the box.

Exercise. A2: Drill

The parts run from one step, through a missing line, a justification column and steps to judge, to short derivations.

(a) A box opens with $\neg(P \wedge Q)$ and ends with $\bot$, a line of the box itself. What may RAA write after it, and what may ¬I?

(b) Line 1 is $\neg\neg(P \to Q)$. What does "$\neg\neg$E 1" write?

(c) To prove $Q \vee R$ by RAA, which statement opens the box?

(d) Let x be a real number. To prove $x \geq 2$ by contradiction, what is assumed?

(e) Line 9 is $P \to Q$, justified "RAA 4–8". What are lines 4 and 8?

(f) Line 6 is $Q \vee R$, justified "$\neg\neg$E 5". What is line 5, and by which rule is such a line usually written?

(g) Fill in the justifications (i)–(iv).

1	$\neg\neg(P \wedge Q)$	
2	$Q \to R$	
3	$P \wedge Q$	(i)
4	Q	(ii)
5	R	(iii)
6	$R \vee S$	(iv)

(h) "Suppose 9 is prime. ... This is impossible. Hence 9 is not prime." Is the last step $\neg$I or RAA?

(i) "Suppose n^2 is not even. ... This is impossible. Hence n^2 is even." Is the last step $\neg$I or RAA?

(j) Which of $\neg\neg$I, $\neg\neg$E, LEM, MT, DS are derived in this chapter by a derivation that uses RAA?

(k) Find the illegal line, and say why it is illegal.

1	$\neg(P \wedge Q)$	
2	$\quad \neg P$	asm, for RAA
3	$\quad\quad P \wedge Q$	asm, for $\neg$I
4	$\quad\quad P$	$\wedge$E 3
5	$\quad\quad \bot$	$\neg$E 4, 2
6	$\quad \neg(P \wedge Q)$	$\neg$I 3–5
7	P	RAA 2–5

(l) A box opens at line 2 with $\neg P$; line 4, a line of that box itself, is $\bot$. Is "5. $\neg\neg P$ — RAA 2–4" correct?

(m) Derive $\neg\neg P,\ P \to Q \vdash Q$.

(n) Derive $\neg P \to P \vdash P$.

(o) Derive $\neg P \to Q,\ \neg Q \vdash P$ in two lines after the premises.

Solution.

(a) RAA writes $P \wedge Q$; $\neg$I writes $\neg\neg(P \wedge Q)$.

(b) $P \to Q$.

(c) $\neg(Q \vee R)$, the negation of the whole goal; not $\neg Q \vee R$.

(d) $x < 2$, that is $\neg(x \geq 2)$; not $x \leq 2$, which $x = 2$ satisfies together with the goal.

(e) Line 4 is the assumption $\neg(P \to Q)$, "asm, for RAA"; line 8 is $\bot$, the last line of the box.

(f) Line 5 is $\neg\neg(Q \vee R)$. Such a line is usually written by $\neg$I, closing a box opened by $\neg(Q \vee R)$ (or it is a premise).

(g) (i) $\neg\neg$E 1; (ii) $\wedge$E 3; (iii) $\to$E 2, 4; (iv) $\vee$I 5.

(h) $\neg$I: the box opens with "9 is prime", and the conclusion is its negation.

(i) RAA: the box opens with the negation of "n^2 is even", and the conclusion is that statement itself.

(j) $\neg\neg$E and LEM; $\neg\neg$I, MT and DS are derived without it.

(k) Line 7: lines 2–5 are not a box (the box opened at line 2 ends at line 6), and RAA needs $\bot$ as a line of the box itself.

(l) No: RAA writes P, the statement whose negation opened the box (clause 3); $\neg\neg P$ is the line $\neg$I writes after the same box.

(m) 3 P, $\neg\neg$E 1; 4 Q, $\to$E 2, 3.

(n) 2 $\neg P$, assumption, for RAA; 3 P, $\to$E 1, 2; 4 $\bot$, $\neg$E 3, 2; 5 P, RAA 2–4.

(o) 3 $\neg\neg P$, MT 1, 2; 4 P, $\neg\neg$E 3.

2.5.11.2 B. Practice

Exercise. B1: With and without the new rule, used later

(a) Derive $\neg\neg P,\ \neg\neg(P \to Q) \vdash \neg\neg Q$ **without** RAA. Then say in one sentence why your derivation does not give Q.

(b) Derive $\neg P \to Q,\ \neg P \to \neg Q \vdash P$, using RAA.

Hint: in (a) the goal is a negation; in (b) it is a letter.

Solution.

(a) The goal $\neg\neg Q$ is a negation, so we open a $\neg$I box on $\neg Q$, with the goal $\bot$. The premises are negations, so each can only clash: $\neg\neg P$ against $\neg P$, and $\neg\neg(P \to Q)$ against $\neg(P \to Q)$. Both of those are negations, so each is itself a subgoal for a $\neg$I box. Make $\neg P$ the subgoal: a box on P. Inside it, make $\neg(P \to Q)$ the subgoal: a box on $P \to Q$, where $\to$E gives Q, which clashes with $\neg Q$. The boxes then close from the inside out.

1	$\neg\neg P$	
2	$\neg\neg(P \to Q)$	
3	$\quad \neg Q$	asm, for $\neg$I
4	$\quad\quad P$	asm, for $\neg$I
5	$\quad\quad\quad P \to Q$	asm, for $\neg$I
6	$\quad\quad\quad Q$	$\to$E 5, 4
7	$\quad\quad\quad \bot$	$\neg$E 6, 3
8	$\quad\quad \neg(P \to Q)$	$\neg$I 5–7
9	$\quad\quad \bot$	$\neg$E 8, 2
10	$\quad \neg P$	$\neg$I 4–9
11	$\quad \bot$	$\neg$E 10, 1
12	$\neg\neg Q$	$\neg$I 3–11

All three boxes close by $\neg$I, so no RAA is used. The derivation does not give Q because its last box opens with $\neg Q$, and $\neg$I can only write the negation of that assumption, $\neg\neg Q$; removing the two negations is $\neg\neg$E, a classical step.

(b) The goal P is a letter, and both premises want $\neg P$, which is nowhere. We are stuck, so we open an RAA box with $\neg P$, and it feeds both premises:

1	$\neg P \to Q$	
2	$\neg P \to \neg Q$	
3	$\quad \neg P$	asm, for RAA
4	$\quad Q$	$\to$E 1, 3
5	$\quad \neg Q$	$\to$E 2, 3
6	$\quad \bot$	$\neg$E 4, 5
7	P	RAA 3–6

In words: "Suppose, for a contradiction, that P fails. Then the premises give both Q and $\neg Q$, which is impossible. Hence P."

Exercise. B2: Back from negated sides, used later

Derive $\neg P \leftrightarrow \neg Q \vdash P \leftrightarrow Q$. Say where RAA enters and why.

Solution. The goal is a biconditional, so by Definition 2.4.1 we need two boxes, one from P to Q and one from Q to P. In the first, the goal Q is a letter and the premise speaks only about $\neg Q$, so we are stuck: we assume $\neg Q$ and argue by RAA. The second box is symmetric.

1	$\neg P \leftrightarrow \neg Q$	
2	$\quad P$	asm, for $\leftrightarrow$I
3	$\quad\quad \neg Q$	asm, for RAA
4	$\quad\quad \neg P$	$\leftrightarrow$E 1, 3
5	$\quad\quad \bot$	$\neg$E 2, 4
6	$\quad Q$	RAA 3–5
7	$\quad Q$	asm, for $\leftrightarrow$I
8	$\quad\quad \neg P$	asm, for RAA
9	$\quad\quad \neg Q$	$\leftrightarrow$E 1, 8
10	$\quad\quad \bot$	$\neg$E 7, 9
11	$\quad P$	RAA 8–10
12	$P \leftrightarrow Q$	$\leftrightarrow$I 2–6, 7–11

RAA enters at lines 6 and 11: each inner goal is a letter, and the premise produces only negations. In the converse, $P \leftrightarrow Q \vdash \neg P \leftrightarrow \neg Q$ (Example (Negating both sides of an iff)), the inner goals are negations, $\neg$I builds them, and no RAA is used.

Exercise. B3: The redundant rule

Show that $\bot$E is derivable from RAA and repetition: given a line $\bot$, derive an arbitrary statement P in one box, without $\bot$E.

Solution. Open an RAA box with $\neg P$ and repeat the $\bot$ inside it:

1	$\bot$	
2	$\quad \neg P$	asm, for RAA
3	$\quad \bot$	R 1
4	P	RAA 2–3

The repetition at line 3 is needed because RAA asks for $\bot$ as a line of the box itself. So every use of $\bot$E can be replaced by these three lines; the book keeps the rule for the reason given in the remark after Theorem 2.5.7.

Exercise. B4: Stuck at a letter

A search for $(\neg P \to Q) \wedge (\neg P \to R),\ \neg(Q \wedge R) \vdash P$ has reached the stage below. *Name the move, then complete it.*

1	$(\neg P \to Q) \wedge (\neg P \to R)$
2	$\neg(Q \wedge R)$
	$\vdots$
10	P

Solution. The goal P is a letter, so no introduction rule applies, and both halves of the first premise wait for $\neg P$, which is nowhere. The move is *when stuck, assume the negation of the goal*: open an RAA box on $\neg P$, and hunt for a clash. The assumption feeds both halves, and $Q \wedge R$ then clashes with the second premise.

1	$(\neg P \to Q) \wedge (\neg P \to R)$	
2	$\neg(Q \wedge R)$	
3	$\quad \neg P$	asm, for RAA
4	$\quad \neg P \to Q$	$\wedge$E 1
5	$\quad Q$	$\to$E 4, 3
6	$\quad \neg P \to R$	$\wedge$E 1
7	$\quad R$	$\to$E 6, 3
8	$\quad Q \wedge R$	$\wedge$I 5, 7
9	$\quad \bot$	$\neg$E 8, 2
10	P	RAA 3–9

The assumption on line 3 is cited twice, at lines 5 and 7; the clash on line 9 is between a line built from it and the second premise.

Exercise. B5: Derive or break, used later

For each claim, derive it with any of the rules of Chapters 1 and 2, or break it by statements about a real number x. In each derivation, mark the classical line (RAA, or $\neg\neg$E or LEM used as a derived rule), if there is one. (A derivation by RAA of $\Gamma \vdash \varphi$ derives $\bot$ from Γ together with $\neg\varphi$: from the counterexample set of the claim, Definition 1.1.3.)

(a) $P \vee Q,\ \neg P \vee R \vdash Q \vee R$.
(b) $\neg(P \wedge Q) \vdash \neg P \wedge \neg Q$.
(c) $P \to Q \vdash (P \leftrightarrow Q) \vee Q$.
(d) $(P \to Q) \vee (P \to R) \vdash P \to (Q \wedge R)$.
(e) $\neg\neg P \vee Q \vdash P \vee Q$.
(f) $\neg(P \leftrightarrow Q) \vdash P \leftrightarrow \neg Q$.
(g) $\neg(P \leftrightarrow Q) \vdash \neg P \wedge Q$.
(h) $(P \wedge Q) \to R \vdash (P \to R) \vee (Q \to R)$.

Solution.

(a) Derive, with no classical line. Split on the first premise; in the case P, split on the second, where the case $\neg P$ closes itself.

1	$P \lor Q$	
2	$\neg P \lor R$	
3	$\quad P$	asm, for $\lor$E
4	$\quad\quad \neg P$	asm, for $\lor$E
5	$\quad\quad \bot$	$\neg$E 3, 4
6	$\quad\quad Q \lor R$	$\bot$E 5
7	$\quad\quad R$	asm, for $\lor$E
8	$\quad\quad Q \lor R$	$\lor$I 7
9	$\quad Q \lor R$	$\lor$E 2, 4–6, 7–8
10	$\quad Q$	asm, for $\lor$E
11	$\quad Q \lor R$	$\lor$I 10
12	$Q \lor R$	$\lor$E 1, 3–9, 10–11

(b) Break. Take P: "$x > 0$", Q: "$x > 5$". At $x = 1$, $P \land Q$ is false, so $\neg(P \land Q)$ is true, and $\neg P$ is false, so $\neg P \land \neg Q$ is false: the counterexample set $\neg(P \land Q)$, $\neg(\neg P \land \neg Q)$ is true. This is the wrong De Morgan law.

(c) Derive; the classical line is line 2, LEM (Proposition 2.5.3 as a derived rule). Each side of the goal breaks on its own: with P: "$x > 5$", Q: "$x > 3$", the premise holds at every x, $P \leftrightarrow Q$ fails at $x = 4$ and Q at $x = 0$. So decide a letter first. The letter is Q: in the case Q the right side is in hand, and in the case $\neg Q$ the premise and the clash with $\neg Q$ build $P \leftrightarrow Q$ in two boxes.

1	$P \to Q$	
2	$Q \lor \neg Q$	LEM
3	$\quad Q$	asm, for $\lor$E
4	$\quad (P \leftrightarrow Q) \lor Q$	$\lor$I 3
5	$\quad \neg Q$	asm, for $\lor$E
6	$\quad\quad P$	asm, for $\leftrightarrow$I
7	$\quad\quad Q$	$\to$E 1, 6
8	$\quad\quad Q$	asm, for $\leftrightarrow$I
9	$\quad\quad \bot$	$\neg$E 8, 5
10	$\quad\quad P$	$\bot$E 9
11	$\quad P \leftrightarrow Q$	$\leftrightarrow$I 6–7, 8–10
12	$\quad (P \leftrightarrow Q) \lor Q$	$\lor$I 11
13	$(P \leftrightarrow Q) \lor Q$	$\lor$E 2, 3–4, 5–12

Deciding P first, as the Idea of Proposition 2.5.6 (b) does, would leave the case $\neg P$ with neither side in hand.

(d) Break. Take P: "$x > 0$", Q: "$x > -1$", R: "$x > 5$". At $x = 1$, $P \to Q$ is true, so the premise is, and $Q \wedge R$ is false while P is true, so $P \to (Q \wedge R)$ is false: the counterexample set is true.

(e) Derive; the classical line is line 3, $\neg\neg$E.

1	$\neg\neg P \vee Q$	
2	$\quad \neg\neg P$	asm, for $\vee$E
3	$\quad P$	$\neg\neg$E 2
4	$\quad P \vee Q$	$\vee$I 3
5	$\quad Q$	asm, for $\vee$E
6	$\quad P \vee Q$	$\vee$I 5
7	$P \vee Q$	$\vee$E 1, 2–4, 5–6

(f) Derive; the classical line is line 21, RAA. The goal is a biconditional, so two boxes. In the first, $\neg Q$ is a negation and $\neg$I builds it; in the second, the goal P is a letter, and an RAA box on $\neg P$ gives the clash.

1	$\neg(P \leftrightarrow Q)$	
2	$\quad P$	asm, for $\leftrightarrow$I
3	$\quad\quad Q$	asm, for $\neg$I
4	$\quad\quad\quad P$	asm, for $\leftrightarrow$I
5	$\quad\quad\quad Q$	R 3
6	$\quad\quad\quad Q$	asm, for $\leftrightarrow$I
7	$\quad\quad\quad P$	R 2
8	$\quad\quad P \leftrightarrow Q$	$\leftrightarrow$I 4–5, 6–7
9	$\quad\quad \bot$	$\neg$E 8, 1
10	$\quad \neg Q$	$\neg$I 3–9
11	$\quad \neg Q$	asm, for $\leftrightarrow$I
12	$\quad\quad \neg P$	asm, for RAA
13	$\quad\quad\quad P$	asm, for $\leftrightarrow$I
14	$\quad\quad\quad \bot$	$\neg$E 13, 12
15	$\quad\quad\quad Q$	$\bot$E 14
16	$\quad\quad\quad Q$	asm, for $\leftrightarrow$I
17	$\quad\quad\quad \bot$	$\neg$E 16, 11
18	$\quad\quad\quad P$	$\bot$E 17
19	$\quad\quad P \leftrightarrow Q$	$\leftrightarrow$I 13–15, 16–18
20	$\quad\quad \bot$	$\neg$E 19, 1
21	$\quad P$	RAA 12–20
22	$P \leftrightarrow \neg Q$	$\leftrightarrow$I 2–10, 11–21

(g) Break. Take P: "$x > 0$", Q: "$x > 5$". At $x = 1$, P is true and Q false, so $P \leftrightarrow Q$ is false and the premise is true; $\neg P$ is false, so $\neg P \wedge Q$ is false.

(h) Derive; the classical line is line 14, RAA. Each side of the goal breaks on its own: with P: "$x > 0$", Q: "$x > 5$", R: "$x > 3$", the premise holds at every x and $P \to R$ fails at $x = 1$; exchanging the statements for P and Q breaks the step to $Q \to R$ at the same x. So assume the negation of the whole goal. Inside, the box on P turns the clash into R, and the box on Q inside it builds $P \wedge Q$.

1	$(P \wedge Q) \to R$	
2	$\quad \neg((P \to R) \vee (Q \to R))$	asm, for RAA
3	$\quad\quad P$	asm, for $\to$I
4	$\quad\quad\quad Q$	asm, for $\to$I
5	$\quad\quad\quad P \wedge Q$	$\wedge$I 3, 4
6	$\quad\quad\quad R$	$\to$E 1, 5
7	$\quad\quad Q \to R$	$\to$I 4–6
8	$\quad\quad (P \to R) \vee (Q \to R)$	$\vee$I 7
9	$\quad\quad \bot$	$\neg$E 8, 2
10	$\quad\quad R$	$\bot$E 9
11	$\quad P \to R$	$\to$I 3–10
12	$\quad (P \to R) \vee (Q \to R)$	$\vee$I 11
13	$\quad \bot$	$\neg$E 12, 2
14	$(P \to R) \vee (Q \to R)$	RAA 2–13

(a) needs no classical line, although its goal is a disjunction, because each case of the premise hands over one side. (c) and (h) are disjunctions whose sides break on their own, the place "When RAA earns its place" points to, and the answers use both classical tools there: (c) decides a letter first with LEM, as Proposition 2.5.6 (b) does, but the letter is Q, a side of the goal, not P; (h) assumes the negation of the goal. In (f) the classical line closes a box whose goal is a letter, P from $\neg Q$, as in Proposition 2.5.1, the other case that paragraph names; the box 2–10 of (f) has no classical line. (b), (d) and (g) are look-alikes. Worked examples of these types: Proposition 2.5.3, Proposition 2.5.5 (d), Proposition 2.5.6 (b), Proposition 2.5.2.

Marking scheme. (16 points) For each claim: 1 for the verdict; 1 for a derivation every line of which checks, with the classical line correctly marked, or for a choice of statements and one value of x at which **every** member of the counterexample set is true. Fatal errors (0 for the claim): a "break" at which a premise fails; a classical line not marked, or a non-classical line marked as classical.

Exercise. B6: A law that sounds false

(a) Show that $(P \to Q) \vee (Q \to P)$ cannot be broken: whatever statements P and Q are, one disjunct is true.

(b) Let n be an integer, P: "n is even", Q: "$n > 5$". Check the statement at $n = 2$ and at $n = 7$, saying which disjunct holds.

(c) The sentence "every even integer is greater than 5, or every integer greater than 5 is even" is false. Why is that no counterexample to (a)?

(d) Show that each disjunct, as a claim from no premises, breaks. Which move of this section does a derivation of $(P \to Q) \vee (Q \to P)$ then call for?

Solution.

(a) If P is false, $P \to Q$ is true, since its hypothesis fails. If P is true, $Q \to P$ is true, since its conclusion holds. Either way one disjunct is true.

(b) At $n = 2$: P is true and Q false, so $Q \to P$ holds (and $P \to Q$ fails). At $n = 7$: P is false and Q true, so $P \to Q$ holds (and $Q \to P$ fails).

(c) The false sentence puts "every" in front of each disjunct separately: (for every n, $P \to Q$) or (for every n, $Q \to P$). The value $n = 2$ refutes the first and $n = 7$ the second. What (a) gives is, for each n, the disjunction itself, as (b) checked at two values. "Every" does not pass inside an "or". This is the point of "Two sequents that sound wrong" in Section 2.1: the mathematical "if" sits under an "every".

(d) $P \to Q$ breaks with P: "$x > 0$", Q: "$x > 5$", at $x = 1$; $Q \to P$ breaks with the same two statements exchanged, at $x = 1$. So $\vee$I from either side is not worth aiming at, and the goal is a disjunction with neither side within reach: the case of "When RAA earns its place", and the move *when stuck, assume the negation of the goal.* (The derivation itself is Exercise C2 of Section 3.1, where the search method finds it; Chapter 4 states, without proof, that every derivation of it uses RAA or one of its equivalents.)

Marking scheme. (10 points) 3: (a), both cases. 2: (b), the disjunct that holds at each value. 3: (c), the scope of "every" named, with a refuting value for each separated disjunct. 2: (d), the two breaks and the move named. Fatal error (0 for (c)): calling the false sentence an instance of (a).

2.5.11.3 C. Going deeper

Exercise. C1: Challenge: Excluded middle does the work of reductio directly

Suppose that, instead of RAA, we had LEM as a rule: $\varphi \vee \neg\varphi$ may be written on any line, for any statement φ.

Given a box that opens with $\neg\varphi$ and reaches $\bot$, derive φ using LEM, $\vee$E, $\bot$E, and the other primitive rules, but neither RAA nor $\neg\neg$E. Write the case $\neg\varphi$ as the given box, extended.

Solution (Hint 1). Write $\varphi \vee \neg\varphi$ just before the given box. What must each of the two case boxes end with?

Solution (Hint 2). The first case box is a single line.

Solution. Write excluded middle for φ just before the given box, and use the box as the second case of $\vee$E, extended by one line:

h	$\varphi \vee \neg\varphi$	LEM
$h+1$	$\quad\vert\ \varphi$	asm, for $\vee$E
i	$\quad\vert\ \neg\varphi$	asm, for $\vee$E
	$\quad\vert\ \vdots$	
j	$\quad\vert\ \bot$	
$j+1$	$\quad\vert\ \varphi$	$\bot$E j
k	φ	$\vee$E h, $h+1$–$h+1$, i–$j+1$

Here $i = h + 2$. The first case is a one-line box that already ends with the goal. The second case is the given box with one more line, $\bot$E on its $\bot$; its lines are unchanged and cite what they cited before, since the box has not moved into or out of any other box. No RAA and no $\neg\neg$E is used.

Exercise. C2: Find the error, used later

A student derives $\neg(\neg P \wedge \neg Q) \vdash P \vee Q$ as follows.

1	$\neg(\neg P \wedge \neg Q)$	
2	$P \vee \neg P$	LEM
3	$\quad\vert\ P$	asm, for $\vee$E
4	$\quad\vert\ P \vee Q$	$\vee$I 3
5	$\quad\vert\ \neg P$	asm, for $\vee$E
6	$\quad\vert\ \bot$	$\neg$E 3, 5
7	$\quad\vert\ Q$	$\bot$E 6
8	$\quad\vert\ P \vee Q$	$\vee$I 7
9	$P \vee Q$	$\vee$E 2, 3–4, 5–8

(a) Is the sequent $\neg(\neg P \wedge \neg Q) \vdash P \vee Q$ derivable?

(b) Does the student's derivation derive it? Name the first line that fails and the rule or clause it breaks.

(c) Show that the illegal step, if it were allowed, would prove too much: use the same move to "derive" $P \vee Q \vdash P$, and break that sequent.

(d) Give a correct derivation.

Solution.

(a) Yes: part (d) below gives a derivation. The conclusion is true, and the argument is still wrong.

(b) No. Line 6 fails: it cites line 3, the assumption of the first case box, lines 3–4, which closed before line 5. By the scope rule of Definition 1.4.2, line 3 is not available at line 6. The form of $\neg$E is right; what is broken is availability. In words, the student used

the assumption of one case inside the other case, where it has not been assumed.

(c) The same move derives $P \vee Q \vdash P$: split on the premise by $\vee$E, with a first box that opens with P and ends there, and a second box that opens with Q and "repeats" P from the first. That sequent breaks: with P false and Q true the premise is true and the conclusion false, so the move can carry a true premise to a false conclusion. So citing a line of a sibling box is illegal in general, not just unlucky here.

(d) Keep lines 1–5, and in the case $\neg P$ we are stuck at the goal Q, so we argue for it by RAA, using the premise:

1	$\neg(\neg P \wedge \neg Q)$	
2	$P \vee \neg P$	LEM
3	$\quad P$	asm, for $\vee$E
4	$\quad P \vee Q$	$\vee$I 3
5	$\quad \neg P$	asm, for $\vee$E
6	$\quad\quad \neg Q$	asm, for RAA
7	$\quad\quad \neg P \wedge \neg Q$	$\wedge$I 5, 6
8	$\quad\quad \bot$	$\neg$E 7, 1
9	$\quad Q$	RAA 6–8
10	$\quad P \vee Q$	$\vee$I 9
11	$P \vee Q$	$\vee$E 2, 3–4, 5–10

Line 7 cites line 5, the assumption of the enclosing case box, which is still open: the legal version of what the student wanted.

Exercise. C3: Exit check: a classical disjunction

Derive $\vdash P \vee (P \to Q)$ using the **primitive** rules only (those of Chapter 1, $\neg$E, $\bot$E, $\neg$I, $\leftrightarrow$I, $\leftrightarrow$E, and RAA), with no derived rule, so neither LEM nor $\neg\neg$E as a single line. Name the rule on every line, and say which line is classical.

Solution. The search. The goal is a disjunction, so we postpone it; there are no premises, so no premise-directed move applies; and neither side is worth aiming at: P false breaks the step to P, and P true with Q false breaks the step to $P \to Q$. We are stuck, so we assume the exact negation of the goal, $\neg(P \vee (P \to Q))$, and hunt for a clash. The assumption can clash only with $P \vee (P \to Q)$, which $\vee$I builds from either side. The side $P \to Q$ is an implication, so we open a box on P with the goal Q. Inside it, $\vee$I gives $P \vee (P \to Q)$ from P, which clashes with the assumption, and $\bot$E writes Q. So $P \to Q$ holds, $\vee$I gives the goal, and that is the second clash.

1	$\neg(P \vee (P \to Q))$	asm, for RAA
2	$\quad P$	asm, for $\to$I
3	$\quad P \vee (P \to Q)$	$\vee$I 2
4	$\quad \bot$	$\neg$E 3, 1
5	$\quad Q$	$\bot$E 4
6	$P \to Q$	$\to$I 2–5
7	$P \vee (P \to Q)$	$\vee$I 6
8	$\bot$	$\neg$E 7, 1
9	$P \vee (P \to Q)$	RAA 1–8

Lines 1 and 2 are assumptions, of the RAA box and of the $\to$I box; every other line names its rule on the right. The classical line is line 9; every other rule used is a rule for a connective. The shape is that of Proposition 2.5.3: there the inner box closed by $\neg$I to give $\neg P$; here it ends with $\bot$E and closes by $\to$I to give $P \to Q$.

A second correct route first derives $P \vee \neg P$ by the eight lines of Proposition 2.5.3, written out, and then splits on it: in the case P, $\vee$I gives the goal; in the case $\neg P$, a box on P clashes with $\neg P$, $\bot$E writes Q, $\to$I gives $P \to Q$, and $\vee$I the goal. Its classical line is the RAA inside the written-out excluded middle.

Marking scheme. (10 points)

- 2: an RAA box opened with $\neg(P \vee (P \to Q))$, the exact negation of the goal.
- 3: an inner box on P that reaches $\bot$ by $\vee$I and $\neg$E against the outer assumption, and then Q by $\bot$E.
- 2: $P \to Q$ by $\to$I, then $P \vee (P \to Q)$ by $\vee$I, and the second clash.
- 2: RAA closes the outer box, and it is named as the classical step.
- 1: every citation is available.

The second route earns the same points, step for step: 2 for the written-out excluded middle opened correctly (its RAA box on $\neg(P \vee \neg P)$); 3 for the box on P in the case $\neg P$, ending with Q by $\bot$E; 2 for $\to$I, the two $\vee$I lines, and the $\vee$E; 2 for naming the RAA inside excluded middle as the classical step; 1 for availability. Fatal errors (0 for the step): opening the outer box with anything but $\neg(P \vee (P \to Q))$; closing the outer box by $\neg$I (that gives $\neg\neg(P \vee (P \to Q))$); a line justified by $\neg\neg$E or LEM taken as a rule, which the item excludes.

Exercise. C4: Challenge: Smaller rule sets

(a) Show that, once RAA is a rule, $\neg$I could be dropped: given a box that opens with φ and reaches $\bot$ as a line of its own, derive $\neg\varphi$ with RAA, $\neg$E and the box's lines, without $\neg$I. Use your method to derive Proposition 2.3.1 without $\neg$I.

(b) With neither $\bot$E nor RAA ($\neg$I allowed), derive $\neg P \vdash P \to \neg Q$.

(c) Exercise B3 derived $\bot$E from RAA and repetition. Does its derivation use $\neg$I, and does your fragment in (a) use $\bot$E? What do (a) and B3 together say about the list of thirteen primitive rules?

(d) Which of the two paradoxes of Section 2.1 survives in a system with neither $\bot$E nor

RAA, by the derivation given there?

Solution.

(a) Replace the old box by this one. Line i: $\neg\neg\varphi$, assumption, for RAA. Line $i+1$: $\neg\varphi$, assumption, for RAA. Line $i+2$: $\bot$, $\neg$E $i+1$, i. Line $i+3$: φ, RAA $i+1$–$i+2$. Then the old box's lines after its assumption, renumbered, each citation of the old assumption now citing line $i+3$; the last of them is $\bot$, a line of the new box. Then $\neg\varphi$, by RAA on the new box: RAA writes ψ from a box opened by $\neg\psi$, here with ψ the statement $\neg\varphi$. Every old citation stays available: the old lines sit in the new box exactly as they sat in the old one, and line $i+3$ is a line of that box.
For modus tollens:

1	$P \to Q$	
2	$\neg Q$	
3	$\quad\neg\neg P$	asm, for RAA
4	$\quad\quad\neg P$	asm, for RAA
5	$\quad\quad\bot$	$\neg$E 4, 3
6	$\quad P$	RAA 4–5
7	$\quad Q$	$\to$E 1, 6
8	$\quad\bot$	$\neg$E 7, 2
9	$\neg P$	RAA 3–8

(b) A box on P, and inside it a box on Q whose $\bot$ comes from the clash of P with the premise:

1	$\neg P$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad Q$	asm, for $\neg$I
4	$\quad\quad\bot$	$\neg$E 2, 1
5	$\quad\neg Q$	$\neg$I 3–4
6	$P \to \neg Q$	$\to$I 2–5

The $\neg$I box never uses its assumption; the fold "Edge cases of not-introduction" of Section 2.2 explains why that is legal: the lines above it, P and $\neg P$, already clash.

(c) No, and no: B3 uses RAA and repetition only, and (a) uses RAA and $\neg$E only. So $\bot$E and $\neg$I can be dropped together, and the remaining eleven rules derive every sequent the thirteen do (replace the lines one at a time, innermost boxes first). The book keeps both for the reason the remark after Theorem 2.5.7 gives for $\bot$E: the primitive rules other than RAA are to form a system of their own.

(d) $Q \vdash P \to Q$: its derivation uses $\to$I and repetition only. The derivation of $\neg P \vdash P \to Q$ given in Section 2.1 uses $\bot$E.

Marking scheme. (10 points) 5: (a), the general replacement (2 for the inner RAA producing φ; 1 for re-citing the old assumption; 1 for closing by RAA with ψ the statement $\neg\varphi$) and 1 for the modus tollens derivation. 2: (b). 2: (c), both answers and the conclusion

about the list. 1: (d). Fatal error (0 for (a)): closing the new box by ¬I, or opening it with φ.

2.5.11.4 R. Review

Exercise. R1

Let n be an integer. Prove that $n^2 + 3n$ is even, naming the rule at each logical step.

Solution. By (E1), n is even or n is odd, and we argue by cases [∨E].

Case n even. Write $n = 2k$ with k an integer. Then $n^2 + 3n = 4k^2 + 6k = 2(2k^2 + 3k)$, which is even.

Case n odd. Write $n = 2k + 1$ with k an integer. Then $n^2 + 3n = (4k^2 + 4k + 1) + (6k + 3) = 2(2k^2 + 5k + 2)$, which is even.

Both cases reach the goal, so $n^2 + 3n$ is even [∨E].

The method is Chapter 1's case split (Definition 1.6.1), used as in Example (The absolute value of a product), here on the parity of n. No rule of this chapter is needed: neither case clashes, and each ends in the goal.

Exercise. R2

Call an integer n **balanced** if n is even exactly when n is a multiple of 3. Write E: "n is even", T: "n is a multiple of 3".

(a) Write "n is balanced" in E and T, and decide it for $n = 4, 6, 7, 9$.

(b) Show that "n is balanced" says the same as "n is both even and a multiple of 3, or neither", by deriving $E \leftrightarrow T \dashv\vdash (E \wedge T) \vee (\neg E \wedge \neg T)$. In which direction does your derivation use a classical line (RAA, ¬¬E or LEM)?

Solution.

(a) $E \leftrightarrow T$: a definition's "if" is read as an iff (Section 2.4, as in Example (A definition used in both directions); the rules are Definition 2.4.1). At 4, E is true and T false: not balanced. At 6, both are true: balanced. At 7, both are false: balanced. At 9, E is false and T true: not balanced.

(b) Left to right; the classical line is line 2, LEM (Proposition 2.5.3 as a derived rule):

1	$E \leftrightarrow T$	
2	$E \vee \neg E$	LEM
3	$\quad E$	asm, for $\vee$E
4	$\quad T$	$\leftrightarrow$E 1, 3
5	$\quad E \wedge T$	$\wedge$I 3, 4
6	$\quad (E \wedge T) \vee (\neg E \wedge \neg T)$	$\vee$I 5
7	$\quad \neg E$	asm, for $\vee$E
8	$\quad\quad T$	asm, for $\neg$I
9	$\quad\quad E$	$\leftrightarrow$E 1, 8
10	$\quad\quad \bot$	$\neg$E 9, 7
11	$\quad \neg T$	$\neg$I 8–10
12	$\quad \neg E \wedge \neg T$	$\wedge$I 7, 11
13	$\quad (E \wedge T) \vee (\neg E \wedge \neg T)$	$\vee$I 12
14	$(E \wedge T) \vee (\neg E \wedge \neg T)$	$\vee$E 2, 3–6, 7–13

Right to left, with no classical line:

1	$(E \wedge T) \vee (\neg E \wedge \neg T)$	
2	$\quad E \wedge T$	asm, for $\vee$E
3	$\quad\quad E$	asm, for $\leftrightarrow$I
4	$\quad\quad T$	$\wedge$E 2
5	$\quad\quad T$	asm, for $\leftrightarrow$I
6	$\quad\quad E$	$\wedge$E 2
7	$\quad E \leftrightarrow T$	$\leftrightarrow$I 3–4, 5–6
8	$\quad \neg E \wedge \neg T$	asm, for $\vee$E
9	$\quad\quad E$	asm, for $\leftrightarrow$I
10	$\quad\quad \neg E$	$\wedge$E 8
11	$\quad\quad \bot$	$\neg$E 9, 10
12	$\quad\quad T$	$\bot$E 11
13	$\quad\quad T$	asm, for $\leftrightarrow$I
14	$\quad\quad \neg T$	$\wedge$E 8
15	$\quad\quad \bot$	$\neg$E 13, 14
16	$\quad\quad E$	$\bot$E 15
17	$\quad E \leftrightarrow T$	$\leftrightarrow$I 9–12, 13–16
18	$E \leftrightarrow T$	$\vee$E 1, 2–7, 8–17

The classical line of these derivations is in the left-to-right direction, the LEM at line 2. There the goal is a disjunction neither of whose sides holds for every balanced n (at $n = 7$ the left fails, at $n = 6$ the right): the case "When RAA earns its place" names.

Exercise. R3

Let n be an integer, P: "n is a multiple of 6", Q: "n is even", so that $P \to Q$ is true. For each statement, say whether it follows from $P \to Q$ by the rules so far, naming the result that gives it, or break the step from $P \to Q$ to it at a value of n: (i) $\neg Q \to \neg P$; (ii) $Q \to P$; (iii) $\neg P \to \neg Q$; (iv) $\neg P \vee Q$; (v) $P \to (P \wedge Q)$; (vi) $P \leftrightarrow Q$. Which of the derivations you cite uses a classical line (RAA, $\neg\neg$E or LEM)?

Solution.

(i) Follows: Proposition 2.3.2 (CP).

(ii) Breaks at $n = 2$: $P \to Q$ is true, Q is true and P false, so $Q \to P$ is false.

(iii) Breaks at $n = 2$: $\neg P$ is true and $\neg Q$ false.

(iv) Follows: Proposition 2.5.6, the direction $P \to Q \vdash \neg P \vee Q$, whose derivation in this section uses LEM. It is the one cited derivation with a classical line.

(v) Follows, with no classical line:

1	$P \to Q$	
2	$\quad P$	asm, for $\to$I
3	$\quad Q$	$\to$E 1, 2
4	$\quad P \wedge Q$	$\wedge$I 2, 3
5	$P \to (P \wedge Q)$	$\to$I 2–4

(vi) Breaks at $n = 2$: P is false and Q true.

The breaks are of the step with this P and Q: at $n = 2$ the premise is true (a multiple of 6 is $6k = 2(3k)$, so $P \to Q$ holds for every n) and the conclusion false, so the form is bad.

Chapter 3

Finding Derivations

Chapters 1 and 2 gave every connective a rule for using it and a rule for proving it, and a derivation, once written, can be checked line by line with no idea at all. Finding one is another matter. An assumption may be any statement whatever, and nothing in the rules says which one to make; yet every Idea so far found its derivation with a handful of moves. *"You know every propositional rule. Given a sequent, how do you find a derivation, or decide to stop looking?"* The chapter answers in three steps. Section 3.1 puts the moves in order as one method, says why the order loses nothing, and watches the method succeed and fail. Section 3.2 shows that the **task** shapes the search, and names the syntactic idea behind every task: an **inconsistent** list, one from which $\bot$ is derivable. Section 3.3 starts from a paragraph of mathematics instead of a sequent. The chapter ends on a question its methods cannot answer: a search that fails proves nothing, so how could anyone show that something is **not** derivable?

What you need. Chapters 1 and 2: the rules for every connective; derivations with boxes and the scope rule (Definition 1.4.2); derived rules (Definition 2.3.1) and Lemma 2.3.3; the classical rules of Section 2.5; and from Section 1.1 the counterexample set (Definition 1.1.3) and the move *try to break it*. Steps about "every" and "some", such as "say $m = 2k$", are used as ordinary reasoning, as in Chapter 2, until Chapter 5 gives them their rules.

Time. About 15 hours on the core track (all of A, one B and one C per section, the exit checks); about 25 hours for everything.

Minimum before *Sets, Relations, and Cardinality*. Chapters 1–6 and 8, core; Sections 7.1–7.2 before its Chapter 3.

3.0.1 Roadmap

- **How to Find a Derivation.** The moves of Chapters 1 and 2 put in order as one method, the reason the order loses nothing, four refinements including the detour rule, searches that succeed and one that fails, and every rule on one card.
- **Four Kinds of Task.** Theorems, equivalences, inconsistent lists and claims whose verdict is withheld: how each task starts its search, and why all four are one question, whether a list derives $\bot$.
- **From Argument to Derivation.** Extracting the argument from a paragraph, naming the facts it leaves unsaid, writing a proof and its derivation side by side, and a library of standard sequents grouped by what they are for.

3.0.2 Named moves

- **Work from both ends.** The method: read the skeleton off the goal, use the resources by their main connectives, keep the classical step for last. (Section 3.1)
- **Find the goal among the resources.** If the current goal is already available, cite it or copy it by R. (Section 3.1)
- **Derive what needs nothing new before opening a box.** Lines written outside the boxes are available in all of them. (Section 3.1)
- **Look first for a clash partner already on a line.** For the goal $\bot$, aim at the statement whose negation you already have. (Section 3.1)
- **No detours.** Never build a line by an introduction rule and take it apart at once by the matching elimination. (Section 3.1)
- **Let the task choose the search.** A theorem starts at the goal, an equivalence needs two derivations, an inconsistency aims at a clash. (Section 3.2)
- **Try to break it first.** Before searching, try to make the counterexample set true. An attempt that is forced into a clash at every turn is usually a derivation read backwards: its forced statements suggest the boxes, and its clash the $\bot$. (Section 3.2)
- **Name the unstated fact.** An argument in prose leaves premises unsaid; write each one down and check it. (Section 3.3)
- **Two columns.** Write the proof beside the derivation it mirrors, and see which lines the prose leaves silent. (Section 3.3)

3.0.3 Summary

The search method reads boxes off the goal and eliminations off the resources, and its goal steps and its classical step never turn a gap that can be filled into one that cannot (the fold of Section 3.1). Postponing a disjunction and the four refinements make it efficient. Every task in Section 3.2 is a question about derivability, and every question about derivability is a question about whether a list is **inconsistent** (Proposition 3.2.2): deriving φ from Γ is showing the counterexample set $\Gamma, \neg\varphi$ inconsistent. An inconsistent list derives every statement (Proposition 3.2.1). Arguments in prose become sequents once their unstated facts are named, and the library of Section 3.3 sorts the standard sequents by what they are for.

Every method of the chapter can show that something **is** derivable, and none can show that something is **not**: a claim that breaks, or a search that stalls, has not been shown underivable. That is Chapter 4's question.

3.0.4 Terms

- **consistent list**: a list from which $\bot$ is not derivable (Definition 3.2.1).
- **counterexample set** (recalled): the premises together with the negated conclusion (Definition 1.1.3).
- **current goal**: the written line below a gap (Definition 3.1.1).
- **gap**: a stretch of a partial derivation still missing (Definition 3.1.1).
- **inconsistent list**: a list from which $\bot$ is derivable (Definition 3.2.1).
- **partial derivation**: a derivation with gaps still to fill (Definition 3.1.1).
- **provably equivalent** (recalled): each derivable from the other (Definition 2.4.2).
- **resources**: the lines available at a gap's current goal (Definition 3.1.1).
- **theorem (of logic)**: a statement derivable from no premises, $\vdash \varphi$ (Definition 1.2.3).

3.1 How to Find a Derivation

We now have every rule for the connectives. Checking a derivation is mechanical, as Definition 1.4.2 makes it: for each line, which rule, which lines, and are they available. Finding one is not, since an assumption may be any statement at all. Yet every Idea in Chapters 1 and 2 found its derivation with the same handful of moves. This section introduces no new rule. It collects those moves into one method, puts them in order, says why the order is right, and then watches the method work: on five searches that succeed, one that fails, and one mathematical goal. It ends with every rule in one place.

3.1.1 The moves so far

Each section since Section 1.2 has added a move, either to the list read off the **goal** (goal-directed: which introduction rule builds the goal's main connective?) or to the list read off the **premises** (premise-directed: which elimination rule uses an available line's main connective?). Here they are together, with the section where each first appeared.

The situation	The move	First met
goal $\varphi \wedge \psi$	two subgoals, φ and ψ, joined by $\wedge$I	Section 1.2
resource $\varphi \wedge \psi$	take it apart by $\wedge$E	Section 1.2
resource $\varphi \to \psi$	fire it on φ by $\to$E; if φ is missing, make it a subgoal	Section 1.3
goal $\varphi \to \psi$	a box opened with φ, with the new goal ψ	Section 1.4
goal $\varphi \vee \psi$	postpone; choose a side only when it is in hand	Section 1.5
resource $\varphi \vee \psi$	two cases, each ending in the current goal	Section 1.6
goal $\bot$	find a clash, or manufacture one; an impossible case closes itself by $\bot$E	Section 2.1
resource $\neg\varphi$	clash it by $\neg$E: make φ a subgoal	Section 2.1
goal $\neg\varphi$	a box opened with φ, with the new goal $\bot$	Section 2.2
a pattern with a name	use the derived rule as one line (MT, DS, ...)	Section 2.3
goal $\varphi \leftrightarrow \psi$	two boxes, one for each direction	Section 2.4
resource $\varphi \leftrightarrow \psi$	cross it by $\leftrightarrow$E once one side is on a line	Section 2.4
stuck	assume the negation of the goal (RAA), or split on $\chi \vee \neg\chi$ (LEM)	Section 2.5

Thirteen moves is a lot to hold in mind at a blank page. The method below turns them into an order of attack. Every assumption a search opens says, in its justification column, which move

opened it: "asm, for →I" is the goal step for →, "asm, for ∨E" the resource step for ∨, "asm, for RAA" step 5. To state it we need a name for a derivation that is only half written.

3.1.2 Partial derivations

Here is the first snapshot of the first search below, for $P \to \neg Q \vdash \neg(P \wedge Q) \wedge (Q \to \neg P)$, drawn as in every section since Section 1.2:

1	$P \to \neg Q$	
	⋮	
7	$\neg(P \wedge Q)$	
	⋮	
13	$Q \to \neg P$	
14	$\neg(P \wedge Q) \wedge (Q \to \neg P)$	∧I 7, 13

Each row ⋮ is a **gap**. The written line just below a gap, here line 7 or line 13, is the gap's **current goal**; it has no justification yet. The lines available at that goal, here the premise for the first gap and the premise and line 7 for the second, are the gap's **resources**. The definition says the same thing exactly.

Definition 3.1.1. Partial Derivation

A **partial derivation** is a list of lines laid out as a derivation (Definition 1.4.2), except that some stretches of lines are still missing. Each missing stretch is a **gap**, drawn as a row ⋮. The line just below a gap is its **current goal**; it is already written, but it has no justification yet, since the lines of the gap are to supply one. The lines available at the current goal, in the sense of Definition 1.4.2, are the gap's **resources**.

In words: every line that has been written is a premise, an assumption, a current goal, or a line justified by a rule from lines and boxes available to it; the boxes are nested and closed as in any derivation; only the gaps are missing. A search starts with a single gap, below the premises and above the conclusion, and each step replaces a gap by lines, boxes, and smaller gaps. It ends when no gap is left, and then the partial derivation is a derivation.

The snapshots in the worked searches below are all partial derivations; a line carries the number it will have in the finished derivation, and a line whose place is not yet fixed carries no number. On paper you would number the lines at the end.

↻ **Second reading** · *A snapshot that is not a partial derivation*

Nonexample. Take a snapshot and leave one of its boxes open at the bottom, so that the conclusion sits inside the box. That is not a partial derivation, even though every line is correct so far: Definition 1.4.2 asks that every box close and that the last line lie outside every box, and a gap cannot repair that, since filling a gap adds lines only above its current goal.

3.1.3 The method

Work from both ends: let the goal choose the boxes, let the premises choose the eliminations, and meet in the middle.

Algorithm. **Searching for a derivation of $\Gamma \vdash \varphi$.**

1. **Set up.** Write the premises Γ at the top and φ at the bottom, with one gap between.
2. **Goal-directed step.** Take the main connective of the current goal (Definition 1.1.2) and set up its introduction rule:
   - $\psi \wedge \chi$: two gaps, with goals ψ and χ (Section 1.2);
   - $\psi \to \chi$: a box opened with ψ; inside it, the new goal is χ (Section 1.4);
   - $\neg\psi$: a box opened with ψ; inside it, the new goal is $\bot$ (Section 2.2);
   - $\psi \leftrightarrow \chi$: two boxes, one from ψ to χ and one from χ to ψ (Section 2.4);
   - $\psi \vee \chi$: **postpone**; do not choose a side yet (Section 1.5).

   Repeat while the current goal is a $\wedge$, $\to$, $\neg$, or $\leftrightarrow$.
3. **Premise-directed step.** Take the main connective of a resource and apply its elimination rule: $\wedge$E freely (Section 1.2); $\to$E and $\leftrightarrow$E when the other statement they need is a resource, or can be made a new goal (Sections 1.3 and 2.4); $\vee$E by opening two boxes, one per disjunct, **each ending in the current goal** (Section 1.6). A derived rule may do the work of several lines (Section 2.3): a resource $\psi \to \chi$ with $\neg\chi$ available gives $\neg\psi$ by MT; a resource $\psi \vee \chi$ with $\neg\psi$ available gives χ by DS.
4. **Clash.** If the current goal is $\bot$, look for a pair χ, $\neg\chi$ among the resources, or make χ the new goal for some resource $\neg\chi$ and manufacture it, starting again at step 2 for that goal. Whatever the current goal, if the resources clash, or a clash with a resource $\neg\chi$ can be manufactured, write $\bot$ and then the goal by $\bot$E: an impossible case closes itself (Section 2.1). When the clash needs a disjunction whose negation is a resource, either side may be chosen to build it: once the disjunction is written, $\bot$ follows.
5. **Stuck.** Use RAA: open a box with the negation of the current goal, whose new goal is $\bot$; if that gives nothing, try the same with a goal further out whose box contains the stalled gap. Or use LEM on a well-chosen statement χ, and split on $\chi \vee \neg\chi$ by $\vee$E. Then return to step 3 (Section 2.5).

When no gap is left, number the lines and check the derivation as Definition 1.4.2 says.

The steps are listed in the order in which to try them, and the order has reasons.

- **Step 2 loses nothing.** If a gap with the goal $\psi \to \chi$ can be filled at all, so can the gap inside a box on ψ with the goal χ.
- **Choosing a side of $\vee$ can leave a goal that does not follow.** From $P \vee Q$, with P true and Q false, the step to Q breaks.
- **Step 5 loses nothing either, and still comes last.** The goal $\bot$ gives step 2 nothing to work with, and a good statement to split on must be chosen.

↻ **Second reading** · *Why the order loses nothing*

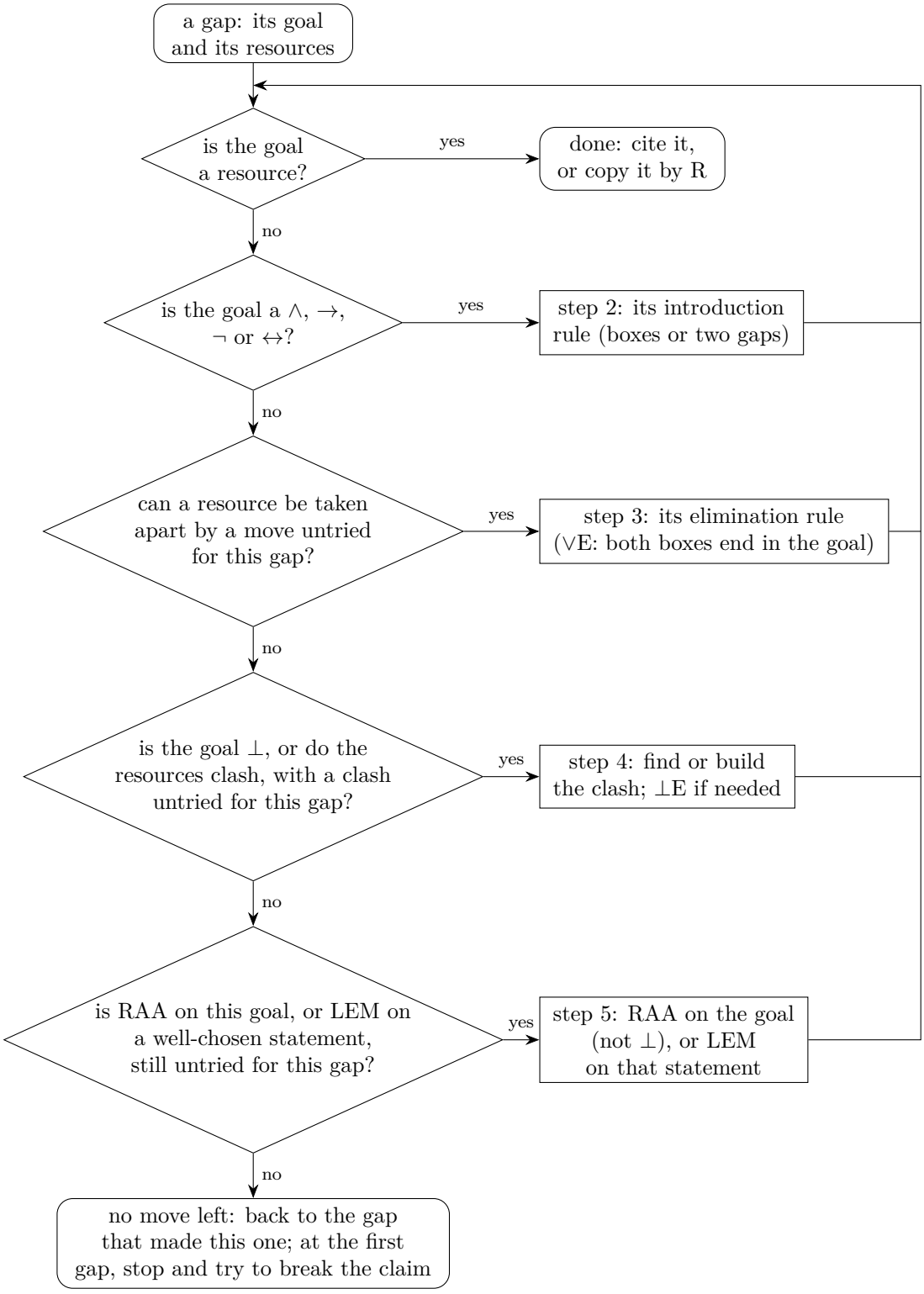
The goal-directed steps lose nothing. Suppose the goal $\psi \to \chi$ is derivable from the resources Γ at all, by some derivation D. Then χ is derivable from Γ together with ψ: add ψ to the premises of D (Definition 1.2.3 asks only that the premises belong to the list), and append the line χ by $\to$E, citing the last line of D and the premise ψ. A derivation from Γ and

ψ can be written inside the box, its premises from Γ cited as available lines and ψ as the box's assumption (the pasting of Lemma 2.3.3). So opening a box on ψ never turns a derivable gap into an underivable one. With $\wedge$E, $\neg$E, or $\leftrightarrow$E as the appended line, the same argument covers $\wedge$, $\neg$, and $\leftrightarrow$. That is why step 2 may be applied without thinking.

Choosing a side of $\vee$ can leave a goal that does not follow. A disjunction can be derivable when neither side follows from the resources. From $P \vee Q$ the goal $Q \vee P$ is derivable (Example (First derivations with "or")), but $P \vee Q$ does not say which side holds, so a last line $\vee$I from Q, or from P, leaves a gap whose goal does not follow from its resources: P true and Q false makes $P \vee Q$ true and Q false. Choosing a side is safe only once the resources decide it, typically inside a case of $\vee$E, or when the disjunction is being manufactured for a clash (step 4) with a resource that is its negation: then the resources derive $\bot$ as soon as they derive the disjunction, and from $\bot$ every statement, either side included, by $\bot$E.

Step 5 also loses nothing, and still comes last. If φ is derivable from the resources, then $\bot$ is derivable from them together with $\neg\varphi$, by one appended $\neg$E; so an RAA box is always safe. A split on $\chi \vee \neg\chi$ is safe too, since each case only adds a resource. Step 5 comes last because the goal $\bot$ gives step 2 nothing to work with, because a good χ must be chosen, and because a derivation that did not need the classical step reads worse with it (the costume of Section 2.5).

The same method as a chart. A move counts only if the gap it makes is not already open above it, with the same goal and resources. When no move is left, go back to the gap that made this one.



Quick check. For each goal, with no premises, name the first rule the method aims for and the box or gaps it sets up.

(a) $\neg(P \wedge \neg P)$

(b) $(P \vee Q) \to (Q \vee P)$

(c) $(P \to Q) \wedge (Q \to (P \to Q))$

(d) $P \vee \neg(P \wedge Q)$

Solution.

(a) ¬I: a box opened with $P \wedge \neg P$, goal $\bot$. (b) $\to$I: a box opened with $P \vee Q$, goal $Q \vee P$; that goal is a disjunction, so it is postponed, and step 3 uses the new resource $P \vee Q$ by $\vee$E. (c) $\wedge$I: two gaps, with goals $P \to Q$ and $Q \to (P \to Q)$, each then a box. (d) Step 2 postpones the disjunction, and with no premises steps 3 and 4 have nothing to use, so the search is at step 5 at once. LEM on P works: in the case P, $\vee$I gives the left side; in the case $\neg P$, a ¬I box on $P \wedge Q$ clashes (its $\wedge$E gives P) and gives $\neg(P \wedge Q)$, and $\vee$I gives the right side.

3.1.4 Splitting the goal

We now run the method. The first search uses only steps 2, 3, and 4. It is the full version of the chapter's main move, *read the skeleton off the goal, then work from both ends*, which comes back twice in the exercises with parts left blank.

Example. Splitting the goal, then using the premise

Find a derivation of $P \to \neg Q \vdash \neg(P \wedge Q) \wedge (Q \to \neg P)$.

Solution. Step 1. The premise on top, the goal at the bottom, one gap. The resources of the gap are the premise alone.

Step 2, first pass. The goal's main connective is $\wedge$, so it splits into two gaps, with goals $\neg(P \wedge Q)$ and $Q \to \neg P$, joined at the end by $\wedge$I:

1	$P \to \neg Q$	
	⋮	
7	$\neg(P \wedge Q)$	
	⋮	
13	$Q \to \neg P$	
14	$\neg(P \wedge Q) \wedge (Q \to \neg P)$	$\wedge$I 7, 13

Step 2, again. Both new goals have a connective, so step 2 applies to each. The goal $\neg(P \wedge Q)$ asks for a ¬I box opened with $P \wedge Q$ and ending in $\bot$. The goal $Q \to \neg P$ asks for a $\to$I box opened with Q; inside it the goal $\neg P$ asks, in turn, for a ¬I box opened with P. Every gap now has the goal $\bot$, and step 2 has nothing more to do:

1	$P \to \neg Q$	
2	$\quad P \wedge Q$	asm, for $\neg$I
	$\quad \vdots$	
6	$\quad \bot$	
7	$\neg(P \wedge Q)$	$\neg$I 2–6
8	$\quad Q$	asm, for $\to$I
9	$\quad\quad P$	asm, for $\neg$I
	$\quad\quad \vdots$	
11	$\quad\quad \bot$	
12	$\quad \neg P$	$\neg$I 9–11
13	$Q \to \neg P$	$\to$I 8–12
14	$\neg(P \wedge Q) \wedge (Q \to \neg P)$	$\wedge$I 7, 13

Steps 3 and 4, first gap. The resources are lines 1 and 2. Step 3 splits line 2 into P and Q by $\wedge$E. Now P is the hypothesis of line 1, so $\to$E gives $\neg Q$. Step 4 finds the pair Q, $\neg Q$.

Steps 3 and 4, second gap. The resources are lines 1, 7, 8, and 9, and line 7 is not needed.

Givens	Goal
$P \to \neg Q$, $\neg(P \wedge Q)$, Q, P	$\bot$
... and $\neg Q$ ($\to$E 1, 9)	$\bot$ ✓ (Q and $\neg Q$ clash)

$\to$E on lines 1 and 9 gives $\neg Q$, which clashes with line 8. Both gaps close:

1	$P \to \neg Q$	
2	$\quad P \wedge Q$	asm, for $\neg$I
3	$\quad P$	$\wedge$E 2
4	$\quad Q$	$\wedge$E 2
5	$\quad \neg Q$	$\to$E 1, 3
6	$\quad \bot$	$\neg$E 4, 5
7	$\neg(P \wedge Q)$	$\neg$I 2–6
8	$\quad Q$	asm, for $\to$I
9	$\quad\quad P$	asm, for $\neg$I
10	$\quad\quad \neg Q$	$\to$E 1, 9
11	$\quad\quad \bot$	$\neg$E 8, 10
12	$\quad \neg P$	$\neg$I 9–11
13	$Q \to \neg P$	$\to$I 8–12
14	$\neg(P \wedge Q) \wedge (Q \to \neg P)$	$\wedge$I 7, 13

What the method did. The goal-directed steps wrote nine of the fourteen lines, every box among them, before the premise was used; the premise-directed steps wrote four, lines 3, 4, 5, and 10. The premise itself was used once in each gap, by the same $\to$E. Line 10 cites line 1 from inside two boxes, which is legal because line 1 is a premise and so is available everywhere below it.

Notice how little thinking the search needed. That is the typical shape of a successful search: **the goal fixes the skeleton, and the premises fill in the flesh.**

Quick check. In the second snapshot of Example (Splitting the goal, then using the premise) above, the second gap has the current goal $\bot$ on line 11, and line 7 is one of its resources, though the gap never uses it. Why is line 7 available there, and why are lines 2–6 not?

Solution. Line 7 lies in no box and above the gap, so it is available, though this gap will not need it. Lines 2–6 are not resources: they lie in a box that has closed.

3.1.5 Cases must end in the goal

When the goal gives nothing, the premises must. The next search is driven by a disjunction among the premises.

Example. Proof by cases, found by the method

Find a derivation of $P \vee Q,\ P \to R,\ Q \to R \vdash R$.

Solution. Step 1. The premises, one gap, the goal:

1	$P \vee Q$
2	$P \to R$
3	$Q \to R$
	⋮
8	R

The goal R is a letter, so step 2 has nothing to do. Step 3 looks at the resources. $\to$E on $P \to R$ needs P, and making P a new goal is hopeless: from $P \vee Q$ we know that one of P, Q holds, not which. The disjunction itself is the resource to use, by $\vee$E, with two boxes that each end in the current goal R:

1	$P \vee Q$	
2	$P \to R$	
3	$Q \to R$	
4	$\quad P$	asm, for $\vee$E
	$\quad$ ⋮	
5	$\quad R$	
6	$\quad Q$	asm, for $\vee$E
	$\quad$ ⋮	
7	$\quad R$	
8	R	$\vee$E 1, 4–5, 6–7

Inside the first box, P is a resource, so $\to$E on line 2 closes the gap at once; the second box is the same with line 3:

1	$P \vee Q$	
2	$P \to R$	
3	$Q \to R$	
4	$\quad P$	asm, for $\vee$E
5	$\quad R$	$\to$E 2, 4
6	$\quad Q$	asm, for $\vee$E
7	$\quad R$	$\to$E 3, 6
8	R	$\vee$E 1, 4–5, 6–7

This is the pattern that Section 1.6 called the skeleton of almost every proof by cases (compare Example (Cases on a hypothesis are two implications), which packages the same two boxes

into one implication). The method found it without being told.

Why the boxes end in the current goal. $\vee$E (Definition 1.6.1) gives, after the two boxes, their common last line and nothing else. A box that ends in anything but the current goal leaves that goal still to be reached, and two boxes that end in different statements cannot be combined at all.

3.1.6 Postponing a disjunction

The third search has a disjunction as its goal. It shows why step 2 postpones one, and how a dead end can name the statement to split on.

Example. Choosing a side too early

Find a derivation of $P \to (Q \vee R) \vdash (P \to Q) \vee (P \to R)$.

Solution. The tempting start. The goal is a disjunction, and it is tempting to choose a side at once, say $P \to Q$, and finish with $\vee$I. The new goal is an implication, so we open a box on P; then $\to$E gives $Q \vee R$, and $\vee$E splits into the cases Q and R, each to end in the current goal Q. The first case ends at once. The second is a dead end:

1	$P \to (Q \vee R)$	
2	$\quad P$	asm, for $\to$I
3	$\quad Q \vee R$	$\to$E 1, 2
4	$\quad\quad Q$	asm, for $\vee$E
5	$\quad\quad R$	asm, for $\vee$E
	$\quad\quad \vdots$	
	$\quad\quad Q$	
	$\quad Q$	$\vee$E
	$P \to Q$	$\to$I
	$(P \to Q) \vee (P \to R)$	$\vee$I

In the case R the resources are $P \to (Q \vee R)$, P, $Q \vee R$, and R, and none of them yields Q: no resource has Q as a conclusion except $Q \vee R$, which is already split, and nothing clashes. Choosing the right side instead fails in the mirror image, in the case Q. In fact neither side follows from the premise alone: with P and Q true and R false the premise is true and $P \to R$ is false, and with P and R true and Q false the premise is true and $P \to Q$ is false.

The diagnosis. Each case of the dead end knew which side it wanted: the case Q wanted $P \to Q$, and the case R wanted $P \to R$. But those cases were opened inside the box of P, too late to choose the side of a disjunction that stands outside it. The side must be decided **before** any box is opened, and the premise does not decide it. Steps 2–4 are used up: we are at step 5.

The dead end also names the statement to split on. If Q holds, the left side is easy. If Q fails, then from P the premise gives $Q \vee R$ with its left side refuted, and DS gives R. So we

use LEM on Q, and each case gets the whole goal:

1	$P \to (Q \vee R)$	
2	$Q \vee \neg Q$	LEM
3	$\quad Q$	asm, for $\vee$E
	$\quad \vdots$	
7	$\quad (P \to Q) \vee (P \to R)$	
8	$\quad \neg Q$	asm, for $\vee$E
	$\quad \vdots$	
13	$\quad (P \to Q) \vee (P \to R)$	
14	$(P \to Q) \vee (P \to R)$	$\vee$E 2, 3–7, 8–13

Now each case decides the side. In the case Q, the goal is $P \to Q$: a box on P whose last line repeats Q. In the case $\neg Q$, the goal is $P \to R$: a box on P, then $Q \vee R$ by $\to$E, then R by Proposition 2.3.4 used as a derived rule, citing $Q \vee R$ and $\neg Q$. Then $\vee$I, on the correct side each time:

1	$P \to (Q \vee R)$	
2	$Q \vee \neg Q$	LEM
3	$\quad Q$	asm, for $\vee$E
4	$\quad\quad P$	asm, for $\to$I
5	$\quad\quad Q$	R 3
6	$\quad P \to Q$	$\to$I 4–5
7	$\quad (P \to Q) \vee (P \to R)$	$\vee$I 6
8	$\quad \neg Q$	asm, for $\vee$E
9	$\quad\quad P$	asm, for $\to$I
10	$\quad\quad Q \vee R$	$\to$E 1, 9
11	$\quad\quad R$	DS 10, 8
12	$\quad P \to R$	$\to$I 9–11
13	$\quad (P \to Q) \vee (P \to R)$	$\vee$I 12
14	$(P \to Q) \vee (P \to R)$	$\vee$E 2, 3–7, 8–13

What the method did. Step 2 postponed the goal, the tempting start broke that rule, and its dead end showed both why the rule exists and what to do instead. Line 5 repeats line 3 because $\to$I needs Q as a line of the box of line 4 itself.

Warning. **Do not choose a disjunct early.** A goal $\psi \vee \chi$ tempts us to pick a side and prove it, and the choice usually fails, as the first attempt above did: neither $P \to Q$ nor

$P \to R$ follows from the premise. Postpone the $\vee$ goal, use the resources first, and choose a side only inside a case that decides it.

3.1.7 When nothing else works

The fourth search reaches step 5 in the other form, RAA, and then needs steps 2 and 4 inside the box.

Example. A search that reaches step 5 at once

Find a derivation of $\neg(P \leftrightarrow Q) \vdash P \vee Q$.

Solution. Steps 1–4. The goal is a disjunction, so step 2 postpones it. The one resource is a negation, and a negation can be used only in a clash (step 4), and here no clash can be found or manufactured. Neither side is in hand: the premise says only that P and Q differ, not which one holds. Steps 2–4 have nothing to offer.

Step 5. We assume the negation of the goal. The new goal is $\bot$, and the gap has one more resource:

1	$\neg(P \leftrightarrow Q)$	
2	$\quad\neg(P \vee Q)$	asm, for RAA
	$\quad\vdots$	
12	$\quad\bot$	
13	$P \vee Q$	RAA 2–12

Step 4 inside the box. The resources are two negations, $\neg(P \leftrightarrow Q)$ and $\neg(P \vee Q)$, and no pair clashes yet. So we choose a statement to manufacture: $P \leftrightarrow Q$, to clash with line 1, or $P \vee Q$, to clash with line 2. The second is the goal we started from, a disjunction of two letters with no side in hand, so aiming at it leads back to the start. So we choose $P \leftrightarrow Q$, and it is a biconditional, so step 2 sets up two boxes:

1	$\neg(P \leftrightarrow Q)$	
2	$\quad\neg(P \vee Q)$	asm, for RAA
3	$\quad\quad P$	asm, for $\leftrightarrow$I
	$\quad\quad\vdots$	
6	$\quad\quad Q$	
7	$\quad\quad Q$	asm, for $\leftrightarrow$I
	$\quad\quad\vdots$	
10	$\quad\quad P$	
11	$\quad P \leftrightarrow Q$	$\leftrightarrow$I 3–6, 7–10
12	$\quad\bot$	$\neg$E 11, 1
13	$P \vee Q$	RAA 2–12

The two small gaps. In the first, the resources are lines 1, 2, and 3, and the goal Q is a letter that no resource produces. But the resources clash once a line is manufactured: $\vee$I builds $P \vee Q$ from line 3, which clashes with line 2, and $\bot$E writes Q (step 4: an impossible case closes itself). The second gap is the mirror image:

1	$\neg(P \leftrightarrow Q)$	
2	$\quad\neg(P \vee Q)$	asm, for RAA
3	$\quad\quad P$	asm, for $\leftrightarrow$I
4	$\quad\quad P \vee Q$	$\vee$I 3
5	$\quad\quad\bot$	$\neg$E 4, 2
6	$\quad\quad Q$	$\bot$E 5
7	$\quad\quad Q$	asm, for $\leftrightarrow$I
8	$\quad\quad P \vee Q$	$\vee$I 7
9	$\quad\quad\bot$	$\neg$E 8, 2
10	$\quad\quad P$	$\bot$E 9
11	$\quad P \leftrightarrow Q$	$\leftrightarrow$I 3–6, 7–10
12	$\quad\bot$	$\neg$E 11, 1
13	$P \vee Q$	RAA 2–12

What the method did. Step 5 came first, because nothing else applied, and after it steps 2 and 4 did the rest: step 4 chose which clash to aim for, step 2 read the two boxes off $P \leftrightarrow Q$, and step 4 closed each box with a manufactured clash and $\bot$E. In words: *suppose neither P nor Q holds; then each of them implies the other, since neither can hold at all; so $P \leftrightarrow Q$, against the premise.* The assumption of line 2 is used twice, at lines 5 and 9.

Second reading · *Other splits that also work*

In Example (Choosing a side too early). LEM on P, the first split one thinks of, also succeeds, but less well: in the case P the premise gives $Q \vee R$, and a second $\vee$E, with two more implication boxes, is needed inside it; the finished derivation has 22 lines against these 14.

In Example (A search that reaches step 5 at once). A split by LEM on P also works, but less well: in the case $\neg P$ the goal side Q is a letter that the resources do not give, and the search is stuck there a second time.

Step 5's advice to try a goal further out is shown on Peirce's law in the fold below.

Second reading · *Step 5 one level out: the search for Peirce's law*

The derivation of Peirce's law (Proposition 2.5.4, in the fold of Section 2.5) adds one detail to step 5.

1	$(P \to Q) \to P$	asm, for $\to$I
	$\quad P$	asm, for $\to$I
	$\quad \vdots$	
	$\quad Q$	
	$P \to Q$	$\to$I
	P	$\to$E
	$((P \to Q) \to P) \to P$	$\to$I

In the box on P with the goal Q, the search stalls: assuming $\neg Q$ gives nothing new, since a clash with $\neg Q$ needs Q itself. One level out, the search assumes $\neg P$:

1	$(P \to Q) \to P$	asm, for $\to$I
2	$\quad \neg P$	asm, for RAA
3	$\quad\quad P$	asm, for $\to$I
	$\quad\quad \vdots$	
5	$\quad\quad Q$	
6	$\quad P \to Q$	$\to$I 3–5
7	$\quad P$	$\to$E 1, 6
8	$\quad \bot$	$\neg$E 7, 2
9	P	RAA 2–8
10	$((P \to Q) \to P) \to P$	$\to$I 1–9

The new resource $\neg P$ is available in every gap inside its box, including the stalled one, where it

clashes with line 3. That is why step 5 says to try a goal further out.

Quick check. In the second snapshot, which two resources of the gap clash, and what does $\bot$E then write?

Solution. Lines 3 (P) and 2 ($\neg P$) give $\bot$ by $\neg$E 3, 2 at line 4, and $\bot$E 4 writes Q, line 5.

3.1.8 Four refinements

The method finds a derivation; these four refinements usually find a shorter or more direct one.

- **Find the goal among the resources.** Before step 2, ask whether the current goal is already available; if so, cite it, or copy it into the current box by R.
- **Derive what needs nothing new before opening a box.** $\wedge$E on a resource, and $\rightarrow$E or $\leftrightarrow$E when both of their lines stand, may be written at once, above the boxes the goal asks for: a line outside the boxes is available inside all of them, so it is written once instead of once per box. ($\vee$E is not such a step: it opens boxes.)
- **Look first for a clash partner already on a line.** When the goal is $\bot$, scan the resources for a negation $\neg\chi$ whose χ is cheap, or for χ and $\neg\chi$ both in reach, before manufacturing anything (step 4).

Example. Lines that need nothing new

Find a derivation of $P \wedge (P \rightarrow Q),\ Q \rightarrow R \vdash (S \rightarrow R) \wedge (T \rightarrow (Q \wedge R))$.

Solution. The search. The goal is a conjunction of two implications, so step 2 will open a box on S and a box on T:

1	$P \wedge (P \rightarrow Q)$
2	$Q \rightarrow R$
	$\vdots$
13	$(S \rightarrow R) \wedge (T \rightarrow (Q \wedge R))$

Before it does, the resources give lines that need nothing new: P and $P \rightarrow Q$ by $\wedge$E, then Q by $\rightarrow$E, and then R by $\rightarrow$E on the second premise. With Q and R outside every box, the two boxes are short:

1	$P \wedge (P \to Q)$	
2	$Q \to R$	
3	P	$\wedge$E 1
4	$P \to Q$	$\wedge$E 1
5	Q	$\to$E 4, 3
6	R	$\to$E 2, 5
7	$\quad S$	asm, for $\to$I
	$\quad \vdots$	
8	$\quad R$	
9	$S \to R$	$\to$I 7–8
10	$\quad T$	asm, for $\to$I
	$\quad \vdots$	
11	$\quad Q \wedge R$	
12	$T \to (Q \wedge R)$	$\to$I 10–11
13	$(S \to R) \wedge (T \to (Q \wedge R))$	$\wedge$I 9, 12

In the box of S the goal R is a resource, copied by R. In the box of T, $\wedge$I joins Q and R:

1	$P \wedge (P \to Q)$	
2	$Q \to R$	
3	P	$\wedge$E 1
4	$P \to Q$	$\wedge$E 1
5	Q	$\to$E 4, 3
6	R	$\to$E 2, 5
7	$\quad S$	asm, for $\to$I
8	$\quad R$	R 6
9	$S \to R$	$\to$I 7–8
10	$\quad T$	asm, for $\to$I
11	$\quad Q \wedge R$	$\wedge$I 5, 6
12	$T \to (Q \wedge R)$	$\to$I 10–11
13	$(S \to R) \wedge (T \to (Q \wedge R))$	$\wedge$I 9, 12

Quick check. In Example (Lines that need nothing new), which refinement does line 8 use, and which do lines 3–6 use?

Solution. Line 8 finds the goal R among the resources (line 6) and copies it into the box of S. Lines 3–6 need nothing new and are written before any box opens, so both boxes can use them.

No detours. Never build a line by an introduction rule and then, at once, take that line apart by the matching elimination. The line the pair produces was within reach without it. This includes ¬I followed by ¬E on the negation it built. A manufactured clash is not a detour: ∨I followed by ¬E against $\neg(\varphi \vee \psi)$ uses the new line with a different connective. The book uses this as advice, not as a theorem: a detour is legal, and cutting one need not shorten a derivation, since an introduced line cited elsewhere stays.

↻ **Second reading** · *Why a detour's line was within reach*

After ∧I joins lines i and j, ∧E on the new line gives back line i or line j. After →I closes a box from φ to ψ, →E on the new implication and a line φ gives ψ; the box's lines, copied with that line in place of the assumption, give ψ directly. After ¬I closes a box from φ to $\bot$, ¬E on the new negation and a line φ gives $\bot$; the box's lines, copied likewise, give $\bot$ directly. After ∨I builds $\varphi \vee \psi$ from a line φ, ∨E on it gives what its first case gives, and the first case's lines, copied with the line φ in place of the assumption, give that directly. Every copied citation stays legal by the scope rule of Definition 1.4.2, since the line put in place of the assumption stands above the copy.

Quick check. Lines 2 and 3 are P and R; line 4 is $P \wedge R$, ∧I 2, 3; line 5 is R, ∧E 4. Which refinement does this break, and what should line 5 cite instead?

Solution. No detours. Lines 4–5 build a conjunction and take it apart at once; line 5's statement is line 3, so cite line 3, or copy it by R where it must lie in the current box.

3.1.9 A search that fails

The method is not a machine that always succeeds. Here it fails, and the failure is instructive.

Try $P \to Q \vdash Q \to P$. Step 2 opens a box on Q with the goal P:

1	$P \to Q$	
2	$\quad Q$	asm, for →I
	$\quad \vdots$	
	$\quad P$	
	$Q \to P$	→I

Every further move goes round in a circle:

Attempt	What it needs	Why that goes round in a circle
step 3: →E on line 1	P	P is the goal
step 4: a clash	a negation among the resources	there is none
step 5: RAA, a box on $\neg P$	P, to clash with $\neg P$	P is the goal
step 5: LEM on P	in the case $\neg P$, the goal P	the same gap, one resource richer and no closer
step 5 further out: RAA on $\neg(Q \to P)$	$Q \to P$, to clash	step 2 opens the same box on Q with the same goal P

Every row returns to one gap, "P from $P \to Q$ and Q". That is affirming the consequent (Section 1.3).

Quick check. Which row of the table would close the gap if affirming the consequent were a rule?

Solution. The first: the mirror image of →E would give P from $P \to Q$ and the resource Q.

But notice what we have **not** shown. A search that fails proves nothing about derivations in general. An assumption may be any statement at all, by the assumption clause of Definition 1.4.2, so there are infinitely many derivations to consider, and a cleverer one might exist. To show that **no** derivation of $Q \to P$ from $P \to Q$ exists, we need a different tool.

We show that something is derivable by exhibiting a derivation; how could we ever show that something is not? Section 1.1's move already breaks Argument B's form, from $P \to Q$ to $Q \to P$: with P false and Q true, the premise is true and the conclusion false. What it cannot yet tell us is that **no derivation** leads from a premise to a conclusion that some choice of statements makes false. Section 3.2 gives this question a sharper form, and the chapter ends on it.

3.1.10 A mathematical goal

The method reads the frame of a proof off the goal whether the statements are letters or facts about numbers. In ordinary mathematics its payoff is that the first and last sentences of a proof, and its division into parts, are fixed before any mathematics is done.

Example. Building a skeleton

Let m and n be integers. Prove that if mn is odd, then m is odd and n is odd.

Solution. The skeleton. The main connective is "if ... then", so step 2 opens a →I box: "Assume mn is odd. ... Hence m is odd and n is odd." The new goal is a conjunction, so step 2 again: two separate proofs, one that m is odd and one that n is odd. Each is a letter as far as logic can see, and the one resource, "mn is odd", cannot be taken apart. So step 3 brings in facts we can compute with: "if m is even then mn is even" is one, and CP (Proposition 2.3.2) turns it into "if mn is not even then m is not even", which the resource

can feed. Then "every integer is even or odd" and DS finish. The givens and the goal are tabulated in the fold below.

The proof. Assume mn is odd. We show that m is odd and that n is odd.

First, if m is even, say $m = 2k$ with k an integer, then $mn = 2(kn)$ is even. By the contrapositive (Proposition 2.3.2), if mn is not even, then m is not even. Now mn is odd, so it is not even, since no integer is both; hence m is not even. Since every integer is even or odd, m is odd, by Proposition 2.3.4.

Exchanging the roles of m and n in the previous paragraph shows that n is odd. Hence m and n are both odd.

What the method did. It fixed the first sentence, the two paragraphs, and the last sentence; the only computation is $mn = 2(kn)$. No step uses RAA: the classical-looking move from "m is not even" to "m is odd" is disjunctive syllogism on the fact that every integer is even or odd.

↻ **Second reading** · *The givens and the goal*

Givens	Goal
the facts: every integer is even or odd, and not both; if m is even, then mn is even; if n is even, then mn is even	if mn is odd, then m is odd and n is odd
… and mn odd ($\to$I)	m odd and n odd
the same	m odd; n odd ($\wedge$I: two goals)
… and mn not even ("not both"); m not even (CP, $\to$E)	m odd ✓ (DS); n odd by the same steps, with m and n exchanged

---

Warning. **Do not assume what you want to prove.** To derive a goal φ, it is tempting to open a box with φ. The box reaches φ at once, but after it closes it gives $\varphi \to \varphi$, not φ, and its assumption is gone. An assumption is useful when the goal asks for it: the hypothesis of an implication (step 2), the negated goal of RAA (step 5), a disjunct of a resource (step 3). Opening the box on the conclusion of an implication is the converse error of Argument B, diagnosed in Section 1.4.

↻ **Second reading** · *Logic permits; it does not choose*

A derivation of $\Gamma \vdash \varphi$ says that whoever accepts every member of Γ is committed to φ. It does not say that φ should now be believed. If φ is plainly false, say $0 = 1$, the derivation is a reason to give up a member of Γ, and it does not say which. A mathematician who derives $0 = 1$ from the hypotheses of a lemma has learned that those hypotheses cannot hold together; Section 3.2 calls such a list inconsistent, and a proof by contradiction is built to reach exactly that.

In the other direction, at every gap the rules permit far more lines than any search writes: $\wedge$E

on every conjunction, $\vee$I with any statement at all. The rules choose none of them. The method chooses, and the method is advice, not logic: two searchers may find two different derivations, and both are correct.

3.1.11 The rules in one place

Here are all the primitive rules of Chapters 1 and 2 on one card.

Rule	Shape	You write	Lean
$\wedge$I	$\varphi, \psi \Longrightarrow \varphi \wedge \psi$	"We show P, and we show Q."	`constructor`
$\wedge$E	$\varphi \wedge \psi \Longrightarrow \varphi$ (or ψ)	"In particular, P."	`h.1` (and `h.2`)
$\to$E	$\varphi \to \psi, \varphi \Longrightarrow \psi$	"Since P implies Q, and P holds, Q."	`exact h hp`
$\to$I	box φ … $\psi \Longrightarrow \varphi \to \psi$	"Assume P. … Hence Q."	`intro hp`
R	φ (available) $\Longrightarrow \varphi$, copied	"As shown above, P."	`exact hp`
$\vee$I	$\varphi \Longrightarrow \varphi \vee \psi$ (or $\psi \vee \varphi$)	"So P or Q holds."	`left` (or `Or.inl hp`); `right`
$\vee$E	$\varphi \vee \psi$, box φ … χ, box ψ … $\chi \Longrightarrow \chi$	"Case 1: P. … so R. Case 2: Q. … so R. In either case, R."	`Or.elim h` (or `cases h`)
$\neg$E	$\varphi, \neg\varphi \Longrightarrow \bot$	"…, which contradicts …"	`exact hn hp` (goal `False`)
$\bot$E	$\bot \Longrightarrow$ any φ	"This case cannot occur." "So there is nothing to prove."	`exfalso`
$\neg$I	box φ … $\bot \Longrightarrow \neg\varphi$	"Suppose P. … This is a contradiction. Hence not P."	`intro hp` (goal `False`)
$\leftrightarrow$I	box φ … ψ, box ψ … $\varphi \Longrightarrow \varphi \leftrightarrow \psi$	"($\Rightarrow$) Suppose P. … so Q. ($\Leftarrow$) Suppose Q. … so P."	`constructor`
$\leftrightarrow$E	$\varphi \leftrightarrow \psi, \varphi \Longrightarrow \psi$ (or with ψ, φ)	"Since P holds if and only if Q does, and P holds, so does Q."	`h.mp hp` / `h.mpr hq`
RAA	box $\neg\varphi$ … $\bot \Longrightarrow \varphi$	"Suppose, for a contradiction, that P fails. … This is impossible. Hence P."	`by_contra h`

This is the reference sheet for Chapters 1 and 2: every primitive rule as its schema, with the label that defines it. As in every schema (Definition 1.2.1), φ, ψ, χ stand for any statements, replaced consistently, and $i, j, k, \ldots$ for line numbers. Each box rule needs the box's last line as a line of that box itself, not inside a further box, and every cited line or box must be available

(Definition 1.4.2).

Conjunction (Definition 1.2.5): ∧I, and the two forms of ∧E.

$$\begin{array}{l|ll} i & \varphi & \\ j & \psi & \\ k & \varphi \wedge \psi & \wedge\text{I } i, j \end{array} \qquad \begin{array}{l|ll} i & \varphi \wedge \psi & \\ k & \varphi & \wedge\text{E } i \end{array} \qquad \begin{array}{l|ll} i & \varphi \wedge \psi & \\ k & \psi & \wedge\text{E } i \end{array}$$

Implication (Definition 1.4.1): →I and →E.

$$\begin{array}{l|ll} i & \boxed{\begin{array}{l|l} & \varphi \\ \hline & \vdots \\ & \psi \end{array}} & \text{asm, for } \to\text{I} \\ k & \varphi \to \psi & \to\text{I } i\text{–}j \end{array} \qquad \begin{array}{l|ll} i & \varphi \to \psi & \\ j & \varphi & \\ k & \psi & \to\text{E } i, j \end{array}$$

Disjunction (Definition 1.6.1): the two forms of ∨I, and ∨E, whose two boxes end in the same statement χ.

$$\begin{array}{l|ll} i & \varphi & \\ k & \varphi \vee \psi & \vee\text{I } i \end{array} \qquad \begin{array}{l|ll} i & \varphi & \\ k & \psi \vee \varphi & \vee\text{I } i \end{array} \qquad \begin{array}{l|ll} i & \varphi \vee \psi & \\ j & \boxed{\begin{array}{l|l} & \varphi \\ \hline & \vdots \\ & \chi \end{array}} & \text{asm, for } \vee\text{E} \\ l & \boxed{\begin{array}{l|l} & \psi \\ \hline & \vdots \\ & \chi \end{array}} & \text{asm, for } \vee\text{E} \\ n & \chi & \vee\text{E } i, j\text{–}k, l\text{–}m \end{array}$$

Contradiction and negation (Definition 2.1.1, Definition 2.2.1): ¬E, ⊥E (which may write **any** statement φ), and ¬I.

$$\begin{array}{l|ll} i & \varphi & \\ j & \neg\varphi & \\ k & \bot & \neg\text{E } i, j \end{array} \qquad \begin{array}{l|ll} i & \bot & \\ k & \varphi & \bot\text{E } i \end{array} \qquad \begin{array}{l|ll} i & \boxed{\begin{array}{l|l} & \varphi \\ \hline & \vdots \\ & \bot \end{array}} & \text{asm, for } \neg\text{I} \\ k & \neg\varphi & \neg\text{I } i\text{–}j \end{array}$$

Biconditional (Definition 2.4.1): ↔I, and ↔E, which works in both directions.

i	|	φ	asm, for $\leftrightarrow$I				
		⋮					
j		ψ		i	$\varphi \leftrightarrow \psi$		
k		ψ	asm, for $\leftrightarrow$I	j	φ		
		⋮		k	ψ	$\leftrightarrow$E i, j	
l		φ					
m	$\varphi \leftrightarrow \psi$		$\leftrightarrow$I i–j, k–l				

i	$\varphi \leftrightarrow \psi$		
j	ψ		
k	φ	$\leftrightarrow$E i, j	

Reductio ad absurdum (Definition 2.5.1), the one classical rule, **and repetition**, clause (D3) of Definition 1.4.2, which copies an available line into the current box.

i	|	$\neg\varphi$	asm, for RAA	i	φ	
		⋮			⋮	
j		$\bot$		k	φ	R i
k	φ		RAA i–j			

Derived rules. Each is a proved sequent, used as one line that cites the lines holding its premises (Definition 2.3.1). The table lists the sequents of Chapters 1 and 2 that searches use most; every other sequent proved there may be used in the same way. Where a rule has no short name, the justification names the result, as "(ex falso)" or "VI".

In a justification	Sequent	Proved in	Our derivation uses RAA
(and commutes)	$P \wedge Q \vdash Q \wedge P$	Proposition 1.2.1	no
(implication is transitive)	$P \to Q,\ Q \to R \vdash P \to R$	Proposition 1.4.1	no
(curry)	$(P \wedge Q) \to R \dashv\vdash P \to (Q \to R)$	Proposition 1.4.2	no
(or commutes)	$P \vee Q \vdash Q \vee P$	Example (First derivations with "or")	no
(or merges)	$P \vee P \vdash P$	Example (First derivations with "or")	no
(and regroups)	$(P \wedge Q) \wedge R \vdash P \wedge (Q \wedge R)$	Proposition 1.2.2	no
(or regroups)	$P \vee (Q \vee R) \vdash (P \vee Q) \vee R$	Example (Regrouping a disjunction)	no
(distributivity)	$P \wedge (Q \vee R) \dashv\vdash (P \wedge Q) \vee (P \wedge R)$	Proposition 1.6.1	no

In a justification	Sequent	Proved in	Our derivation uses RAA
(cases on a hypothesis)	$(P \to R) \wedge (Q \to R) \dashv\vdash (P \vee Q) \to R$	Example (Cases on a hypothesis are two implications); the converse Example (A disjunctive hypothesis splits into two implications)	no
CD (constructive dilemma)	$(P \to R) \wedge (Q \to S),\ P \vee Q \vdash R \vee S$	Example (Constructive dilemma)	no
(destructive dilemma)	$(P \to R) \wedge (Q \to S),\ \neg R \vee \neg S \vdash \neg P \vee \neg Q$	Exercise (B4: Dilemmas as derived rules, used later) (b)	no
(*modus ponendo tollens*)	$\neg(P \wedge Q) \vdash P \to \neg Q$	Exercise (B4: Dilemmas as derived rules, used later) (c)	no
(ex falso)	$P,\ \neg P \vdash Q$	Proposition 2.1.1	no
VI (vacuous implication)	$\neg P \vdash P \to Q$	Proposition 2.1.2	no
MT	$P \to Q,\ \neg Q \vdash \neg P$	Proposition 2.3.1	no
CP	$P \to Q \vdash \neg Q \to \neg P$	Proposition 2.3.2	no
DS	$P \vee Q,\ \neg P \vdash Q$	Proposition 2.3.4	no
(DS, other disjunct)	$P \vee Q,\ \neg Q \vdash P$	Exercise (B1: The other disjunctive syllogism, used later) (a)	no
(iff symmetric)	$P \leftrightarrow Q \vdash Q \leftrightarrow P$	Example (An iff read both ways)	no
(iff as two implications)	$P \leftrightarrow Q \dashv\vdash (P \to Q) \wedge (Q \to P)$	Proposition 2.4.1	no
(iff of negations)	$P \leftrightarrow Q \vdash \neg P \leftrightarrow \neg Q$, and its converse	Example (Negating both sides of an iff); the converse Exercise (B2: Back from negated sides, used later)	the converse only
(equivalence cycle)	$P \to Q,\ Q \to R,\ R \to P \vdash P \leftrightarrow Q$	Proposition 2.4.6	no
¬¬I	$P \vdash \neg\neg P$	Proposition 2.5.1 (a), first in Example (Two negations from one premise)	no
¬¬E	$\neg\neg P \vdash P$	Proposition 2.5.1 (b)	yes
(CP, converse)	$\neg Q \to \neg P \vdash P \to Q$	Proposition 2.5.2	yes
LEM	$\vdash P \vee \neg P$	Proposition 2.5.3	yes
(Peirce)	$\vdash ((P \to Q) \to P) \to P$	Proposition 2.5.4	yes

In a justification	Sequent	Proved in	Our derivation uses RAA
DM	the four laws of Proposition 2.5.5	Proposition 2.5.5	(d) only
(implication as disjunction)	$\neg P \vee Q \vdash P \to Q$, and its converse	Proposition 2.5.6	the converse only, through LEM

Four rows are proved in exercises, not in the text: the other form of DS, the destructive dilemma, *modus ponendo tollens*, and the converse of the iff of negations. Exercises may use them as rules; the text's own proofs do not.

Each rule is also a move: the table of moves at the start of this section is the decision chart of Section 1.6, grown to every connective, read by the goal for the introduction rules and by an available line for the elimination rules; the chart after the method's steps draws their order.

Warning. **Cite only what is already proved.** A derived rule is an abbreviation for its derivation, so it may be used only after that derivation exists, and never inside it. For example, the second route to Proposition 2.5.5 (d) splits on excluded middle; a proof of excluded middle that used De Morgan's law (d) along that route would be circular. In this book the check is simple: cite only labels that appear earlier.

3.1.12 Exercises

3.1.12.1 A. Check your understanding

Exercise. A1

For (a)–(d), name the goal's main connective, the first rule the method aims for, and the first sentence of a proof. For (e)–(f), name the elimination rule the premise invites, and what it needs.

(a) For real x: if $x^3 < 0$ then $x < 0$.
(b) There is no integer n with $2n = 7$.
(c) For real x and y: $|xy| = |x||y|$ and $|x + y| \leq |x| + |y|$.
(d) For an integer n: n is odd if and only if n^2 is odd.
(e) The premise $(P \wedge Q) \to R$, with the goal R and the resources P and Q.
(f) The premise $\neg(P \vee Q)$, with the goal $\bot$ and the resource Q.

Solution.

(a) $\to$; $\to$I: "Assume $x^3 < 0$." (Starting "Suppose $x \geq 0$" inside that frame is an RAA box on the goal $x < 0$, not the contrapositive; a proof through the contrapositive replaces the frame, proving "if $x \geq 0$ then $x^3 \geq 0$" and then applying CP.)
(b) $\neg$, since the claim is "not: there is an integer n with $2n = 7$"; $\neg$I: "Suppose n is an integer with $2n = 7$."
(c) $\wedge$; $\wedge$I: "We prove the two statements in turn."

(d) $\leftrightarrow$; $\leftrightarrow$I: "($\Rightarrow$) Assume n is odd."

(e) $\rightarrow$E, which needs the hypothesis $P \wedge Q$: make it a new goal, reached by $\wedge$I from P and Q; then $\rightarrow$E gives R.

(f) $\neg$E, which needs $P \vee Q$ beside $\neg(P \vee Q)$: make $P \vee Q$ the new goal (step 4), reached by $\vee$I from Q.

Exercise. A2: Drill

The parts run from one step of the method, through available lines and steps to judge, to short derivations.

(a) The goal is $\neg(P \vee Q)$. What does step 2 set up?

(b) The goal is $(P \rightarrow Q) \wedge \neg R$. Which gaps does step 2 leave?

(c) The goal is $\bot$, and the resources are $\neg(P \wedge Q)$, P, Q. What are the next two lines?

(d) The goal is $P \vee Q$, and the one resource is $\neg P \rightarrow Q$. Which step does the work, and how?

(e) In a partial derivation the current goal is line 9, inside the boxes opened at lines 3 and 6. Line 1 is a premise, and lines 4–5 form a box inside box 3 that has closed. Which of lines 1, 3, 4, 6 are resources?

(f) Which step of the method opens an RAA box, and when?

(g) True or false: a failed search shows that the sequent is not derivable.

(h) The goal is R, and the resources are $P \vee Q$, $P \rightarrow R$, $Q \rightarrow S$. Should the search be expected to succeed?

(i) Lines 3 and 4 are $Q \rightarrow R$ (a premise) and Q. A searcher writes 5 Q, assumption; 6 R, $\rightarrow$E 3, 5; 7 $Q \rightarrow R$, $\rightarrow$I 5–6; 8 R, $\rightarrow$E 7, 4. Which refinements of the section does this break, and what single line replaces lines 5–8?

(j) Find a derivation of $P \vee Q,\ \neg P \vdash Q \vee R$.

(k) Find a derivation of $(P \vee Q) \rightarrow R \vdash Q \rightarrow R$.

(l) Find a derivation of $P \rightarrow (Q \rightarrow R),\ P \wedge \neg R \vdash \neg Q$.

Solution.

(a) A box opened with $P \vee Q$, with the goal $\bot$.

(b) Two boxes: one on P with the goal Q, and one on R with the goal $\bot$.

(c) $P \wedge Q$ by $\wedge$I, then $\bot$ by $\neg$E.

(d) Step 2 postpones the goal. Step 3 needs $\neg P$, which is not a resource, and making it a goal stalls. So step 5: LEM on P, whose negation the resource waits for.

(e) 1, 3 and 6; not 4.

(f) Step 5, when steps 2–4 give nothing.

(g) False: it shows only that this search failed.

(h) No. P false, Q true, R false, S true makes every resource true and R false, so the claim breaks. In $\vee$E, the case Q gives only S.

(i) Two refinements. Lines 7–8 are a detour ($\rightarrow$I, then $\rightarrow$E on its own implication). Line 7 rebuilds a resource, and *find the goal among the resources* would have cited line 3.

Replace lines 5–8 by 5 R, $\to$E 3, 4.

(j) 3 Q, DS 1, 2; 4 $Q \lor R$, $\lor$I 3.

(k) 2 Q, assumption, for $\to$I; 3 $P \lor Q$, $\lor$I 2; 4 R, $\to$E 1, 3; 5 $Q \to R$, $\to$I 2–4.

(l) 3 P, $\land$E 2; 4 $\neg R$, $\land$E 2; 5 $Q \to R$, $\to$E 1, 3; 6 $\neg Q$, MT 5, 4.

3.1.12.2 B. Practice

Exercise. B1: Three searches

For each sequent, name the first goal-directed step, then find a derivation.

(a) $P \land (Q \to R) \vdash (P \to Q) \to R$.

(b) $\neg(P \to Q) \vdash P \land \neg Q$. Which conjunct does your derivation obtain by RAA, and which by $\neg$I alone?

(c) $\neg P \to Q \vdash P \lor Q$.

Hint: in (c), look at what the premise is waiting for.

Solution.

(a) The goal is an implication: a box on $P \to Q$ with the goal R, a letter. Step 3 splits the premise; its half $Q \to R$ gives R once we have Q, and the box's assumption gives Q once we have P, which the premise's other half supplies. So the chain runs backwards from R to Q to P, and forwards it reads:

1	$P \land (Q \to R)$	
2	$\quad P \to Q$	asm, for $\to$I
3	$\quad P$	$\land$E 1
4	$\quad Q$	$\to$E 2, 3
5	$\quad Q \to R$	$\land$E 1
6	$\quad R$	$\to$E 5, 4
7	$(P \to Q) \to R$	$\to$I 2–6

(b) The goal is a conjunction: two gaps, with goals P and $\neg Q$. The premise can be used only by $\neg$E, against $P \to Q$, so in each gap we aim to manufacture $P \to Q$. The goal $\neg Q$ is a negation: a box on Q, where $P \to Q$ follows by a box that repeats Q, and it clashes with the premise. The goal P is a letter, and the resources give nothing, so step 5: RAA on $\neg P$. Then $P \to Q$ follows by vacuous implication (Proposition 2.1.2, written VI), and it clashes with the premise. **The conjunct P is the one our derivation obtains by** RAA; $\neg Q$ needs only $\neg$I. (That no derivation of P from the premise avoids RAA is true, but we do not prove it.)

1	$\neg(P \to Q)$	
2	$\quad \neg P$	asm, for RAA
3	$\quad P \to Q$	VI 2
4	$\quad \bot$	¬E 3, 1
5	P	RAA 2–4
6	$\quad Q$	asm, for ¬I
7	$\quad\quad P$	asm, for →I
8	$\quad\quad Q$	R 6
9	$\quad P \to Q$	→I 7–8
10	$\quad \bot$	¬E 9, 1
11	$\neg Q$	¬I 6–10
12	$P \wedge \neg Q$	∧I 5, 11

Written with primitive rules only, line 3 becomes the four lines of a box on P that clashes with $\neg P$, writes Q by $\bot$E, and closes by →I; the derivation then has 15 lines instead of 12.

(c) The goal is a disjunction, so step 2 postpones it; the premise is waiting for $\neg P$, which is not a resource, and nothing clashes. Step 5, LEM on P, the statement the premise is waiting for: in the case P, ∨I gives the left side; in the case $\neg P$, the premise gives Q and ∨I the right side.

1	$\neg P \to Q$	
2	$P \vee \neg P$	LEM
3	$\quad P$	asm, for ∨E
4	$\quad P \vee Q$	∨I 3
5	$\quad \neg P$	asm, for ∨E
6	$\quad Q$	→E 1, 5
7	$\quad P \vee Q$	∨I 6
8	$P \vee Q$	∨E 2, 3–4, 5–7

The derivation has 8 lines. An RAA box on $\neg(P \vee Q)$ also works, with two manufactured clashes inside it.

Exercise. B2: The last step

A search for $P \to Q,\ R \to \neg Q \vdash P \to \neg R$ has reached the stage below; only the last line is missing. *Name the move, then complete it.*

1	$P \to Q$	
2	$R \to \neg Q$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q$	$\to$E 1, 3
5	$\quad\quad R$	asm, for $\neg$I
6	$\quad\quad \neg Q$	$\to$E 2, 5
7	$\quad\quad \bot$	$\neg$E 4, 6
8	$\quad \neg R$	$\neg$I 5–7
9	**?**	?

Solution. The move is the chapter's main one: the skeleton was read off the goal, and the gap was filled from both ends. The goal $P \to \neg R$ is an implication, so the outer box opened with P (line 3); its new goal $\neg R$ is a negation, so the inner box opened with R (line 5). The premises then met the goal in the middle: $\to$E at lines 4 and 6 gave Q and $\neg Q$, and the clash at line 7. The inner box was closed at line 8. What remains is to close the outer box on its goal, by $\to$I 3–8:

1	$P \to Q$	
2	$R \to \neg Q$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q$	$\to$E 1, 3
5	$\quad\quad R$	asm, for $\neg$I
6	$\quad\quad \neg Q$	$\to$E 2, 5
7	$\quad\quad \bot$	$\neg$E 4, 6
8	$\quad \neg R$	$\neg$I 5–7
9	$P \to \neg R$	$\to$I 3–8

The box runs from its assumption P, line 3, to its last line $\neg R$, line 8, which lies in the outer box itself and not in the inner one, as $\to$I requires.

Exercise. B3: Two boxes and a chain

Find a derivation of $P \to (Q \to R),\ P \to (R \to S) \vdash P \to (Q \to S)$, and describe the search in three sentences, naming its goal-directed and premise-directed steps.

Solution.

1	$P \to (Q \to R)$	
2	$P \to (R \to S)$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q \to R$	$\to$E 1, 3
5	$\quad R \to S$	$\to$E 2, 3
6	$\quad\quad Q$	asm, for $\to$I
7	$\quad\quad R$	$\to$E 4, 6
8	$\quad\quad S$	$\to$E 5, 7
9	$\quad Q \to S$	$\to$I 6–8
10	$P \to (Q \to S)$	$\to$I 3–9

The search: the goal is an implication, so a goal-directed step opens a box on P with the goal $Q \to S$, and a second goal-directed step opens a box on Q with the goal S. With P a resource, premise-directed steps fire both premises by $\to$E, giving $Q \to R$ and $R \to S$. Then a backward chain from the goal S (it needs R, which needs Q, a resource) is written forwards as two more $\to$E lines, and the two boxes close in the reverse order of opening. (Lines 4 and 5 may equally be written inside the box of Q.)

Exercise. B4: Cut the detours

The derivation below of $P \wedge Q,\ Q \to R \vdash R \wedge P$ is correct but contains detours.

1	$P \wedge Q$	
2	$Q \to R$	
3	Q	$\wedge$E 1
4	$\quad Q$	asm, for $\to$I
5	$\quad R$	$\to$E 2, 4
6	$Q \to R$	$\to$I 4–5
7	R	$\to$E 6, 3
8	P	$\wedge$E 1
9	$R \wedge P$	$\wedge$I 7, 8
10	R	$\wedge$E 9
11	$R \wedge P$	$\wedge$I 10, 8

(a) Find each pair of lines in which a line is built by an introduction rule and taken apart at once by the matching elimination.

(b) Which line rebuilds a premise?

(c) Rewrite the derivation without the detours.

(d) To cut the →I/→E detour, the box is copied with a line in place of its assumption: which line, and what does the copy give?

Solution.

(a) Lines 6–7: →I, then →E on line 6. Lines 9–10: ∧I, then ∧E on line 9. (Line 11 then rebuilds line 9.)

(b) Line 6 is premise 2, rebuilt by a box. *Find the goal among the resources* would have cited line 2.

(c) Six lines instead of eleven:

1	$P \wedge Q$	
2	$Q \to R$	
3	Q	∧E 1
4	R	→E 2, 3
5	P	∧E 1
6	$R \wedge P$	∧I 4, 5

(d) Line 3 takes the place of the assumption on line 4. The box's one other line, "R, →E 2, 4", becomes "R, →E 2, 3", which is line 4 of (c).

3.1.12.3 C. Going deeper

Exercise. C1: The second half

A search for $P \vee (Q \wedge R) \vdash (P \vee Q) \wedge (P \vee R)$ has used the premise by ∨E and finished the first case; the second case is blank. *Name the move, then complete it.*

1	$P \vee (Q \wedge R)$	
2	$\quad P$	asm, for ∨E
3	$\quad P \vee Q$	∨I 2
4	$\quad P \vee R$	∨I 2
5	$\quad (P \vee Q) \wedge (P \vee R)$	∧I 3, 4
6	$\quad Q \wedge R$	asm, for ∨E
	$\quad \vdots$	
11	$\quad (P \vee Q) \wedge (P \vee R)$	
12	$(P \vee Q) \wedge (P \vee R)$	∨E 1, 2–5, 6–11

Solution (Hint 1). The current goal of the blank box is a conjunction. What does step 2 do with it, and what does step 3 do with the box's assumption?

Solution (Hint 2). Line 7 is Q, by $\wedge$E from line 6.

Solution. The move is the chapter's main one again: read the skeleton off the goal, then work from both ends. The current goal $(P \vee Q) \wedge (P \vee R)$ is a conjunction, so step 2 splits it into the goals $P \vee Q$ and $P \vee R$, to be joined by $\wedge$I. Both are disjunctions, so they are postponed, and step 3 turns to the resource $Q \wedge R$: $\wedge$E gives Q and R. Now each disjunction has a side that is a resource, so $\vee$I is safe.

1	$P \vee (Q \wedge R)$	
2	$\quad P$	asm, for $\vee$E
3	$\quad P \vee Q$	$\vee$I 2
4	$\quad P \vee R$	$\vee$I 2
5	$\quad (P \vee Q) \wedge (P \vee R)$	$\wedge$I 3, 4
6	$\quad Q \wedge R$	asm, for $\vee$E
7	$\quad Q$	$\wedge$E 6
8	$\quad R$	$\wedge$E 6
9	$\quad P \vee Q$	$\vee$I 7
10	$\quad P \vee R$	$\vee$I 8
11	$\quad (P \vee Q) \wedge (P \vee R)$	$\wedge$I 9, 10
12	$(P \vee Q) \wedge (P \vee R)$	$\vee$E 1, 2–5, 6–11

Both boxes end in the current goal of the $\vee$E, as that rule requires. Here $\vee$I adds P on the left of Q and of R; in the first case it added Q and R on the right of P. Citing line 2 inside the second box would be the sibling-box error of Exercise (C2: Find the error, used later).

Exercise. C2: A disjunction from nothing

(a) Derive $\vdash (P \to Q) \vee (Q \to P)$.

(b) In a paragraph, say which step of your derivation uses RAA or LEM, and why the method cannot avoid step 5 on this goal.

Solution (Hint 1). The goal is a disjunction and there are no premises. Which steps of the method does that leave?

Solution (Hint 2). Step 5 has two forms. Which statement would make each side of the goal easy, if you knew whether it held?

Solution.

(a) LEM on Q. In the case Q, the implication $P \to Q$ holds by a box that repeats Q. In the case $\neg Q$, the implication $Q \to P$ holds by a box on Q that clashes with $\neg Q$, followed by $\bot$E.

1	$Q \vee \neg Q$	LEM
2	$\quad Q$	asm, for $\vee$E
3	$\quad\quad P$	asm, for $\to$I
4	$\quad\quad Q$	R 2
5	$\quad P \to Q$	$\to$I 3–4
6	$\quad (P \to Q) \vee (Q \to P)$	$\vee$I 5
7	$\quad \neg Q$	asm, for $\vee$E
8	$\quad\quad Q$	asm, for $\to$I
9	$\quad\quad \bot$	$\neg$E 8, 7
10	$\quad\quad P$	$\bot$E 9
11	$\quad Q \to P$	$\to$I 8–10
12	$\quad (P \to Q) \vee (Q \to P)$	$\vee$I 11
13	$(P \to Q) \vee (Q \to P)$	$\vee$E 1, 2–6, 7–12

LEM on P works equally well, with the roles of the two sides exchanged.

(b) LEM is used at line 1; RAA appears only inside the derivation of LEM (Proposition 2.5.3) that line 1 abbreviates. The method cannot avoid step 5 here. Step 2 postpones the $\vee$ goal; with no premises, step 3 has nothing to eliminate and step 4 nothing to clash; and $\vee$I cannot come first, because neither side follows from no premises (with P true and Q false, $P \to Q$ is false; with P false and Q true, $Q \to P$ is false; Chapter 4 turns such choices into proofs). So a resource must be created from nothing, and only step 5 does that. This shows that the **method** needs step 5; that every derivation of the sequent uses RAA or an equivalent is a stronger claim, which we do not prove.

Marking scheme. (10 points)

- 3: a correct split by LEM (on Q or on P), with two case boxes combined by $\vee$E, or a correct RAA box on the negated goal.
- 2: in one case, the implication by a box that repeats the case assumption.
- 2: in the other case, the implication by a box that clashes with the case assumption, then $\bot$E, or by VI (Proposition 2.1.2).
- 1: $\vee$I on the correct side in each case, every citation available.
- 2: the paragraph names the LEM line (or the RAA line) and explains why no other step of the method applies.

Fatal errors (0 for the step): choosing a side of the goal before the split; citing one case's assumption in the other case. A correct derivation by RAA on the negated goal earns full marks: its 2 + 2 + 1 points go to the two implications derived inside the RAA box (one by

a box that repeats an assumption, one by a clash and $\bot$E, or by VI (Proposition 2.1.2) once $\neg Q$ or $\neg P$ has been derived) and to the $\vee$I lines that clash with the RAA assumption.

Exercise. C3: Exit check: search

Find a derivation of $\neg(\neg P \wedge Q) \vdash P \vee \neg Q$ by the method of this section. Write one sentence for each step of the method you take, and point to the classical step of your derivation.

Solution. The search.

1. *Step 1.* The premise $\neg(\neg P \wedge Q)$ on top, the goal $P \vee \neg Q$ at the bottom, one gap.
2. *Step 2.* The goal is a disjunction, so it is postponed: the premise does not say which side holds.
3. *Step 3.* Nothing can be eliminated. The premise is a negation, so it can be used only in a clash (step 4), against $\neg P \wedge Q$, and we do not have it.
4. *Step 4.* The goal is not $\bot$, and nothing clashes. We are stuck.
5. *Step 5.* We assume the negation of the goal, $\neg(P \vee \neg Q)$, with the new goal $\bot$.
6. *Step 4, inside the box.* The new assumption clashes with $P \vee \neg Q$, which $\vee$I builds from either side. The side P is a letter that nothing gives; the side $\neg Q$ is a negation, which step 2 can build. So $\neg Q$ becomes the goal: a box on Q, with the goal $\bot$. Choosing a side is safe here, since the disjunction is manufactured for a clash with a resource that is its negation (step 4 allows either side when the disjunction is built for a clash with its own negation).
7. *Step 4, in the box on Q.* The resources are the premise, the assumption of step 5, and Q. We aim at the premise: $\neg P \wedge Q$ becomes the goal, and step 2 splits it. Its half Q is a resource; its half $\neg P$ is a negation: a box on P, with the goal $\bot$.
8. *Step 4, in the box on P.* $\vee$I builds $P \vee \neg Q$ from P, which clashes with the assumption of step 5. Now the boxes close from the inside out: $\neg P$, then $\neg P \wedge Q$, which clashes with the premise; then $\neg Q$, then $P \vee \neg Q$, which clashes with the assumption of step 5; and RAA closes the outer box.

1	$\neg(\neg P \wedge Q)$	
2	$\quad \neg(P \vee \neg Q)$	asm, for RAA
3	$\quad\quad Q$	asm, for $\neg$I
4	$\quad\quad\quad P$	asm, for $\neg$I
5	$\quad\quad\quad P \vee \neg Q$	$\vee$I 4
6	$\quad\quad\quad \bot$	$\neg$E 5, 2
7	$\quad\quad \neg P$	$\neg$I 4–6
8	$\quad\quad \neg P \wedge Q$	$\wedge$I 7, 3
9	$\quad\quad \bot$	$\neg$E 8, 1
10	$\quad \neg Q$	$\neg$I 3–9
11	$\quad P \vee \neg Q$	$\vee$I 10
12	$\quad \bot$	$\neg$E 11, 2
13	$P \vee \neg Q$	RAA 2–12

The classical step is line 13, the RAA that closes the box of line 2; every other box closes by $\neg$I. The assumption of line 2 is used twice, at lines 6 and 12, just as in Example (A search that reaches step 5 at once).

A second route. Step 5 may take its other form: LEM on P. In the case P, $\vee$I gives the goal. In the case $\neg P$, the goal's side $\neg Q$ is a negation: a box on Q builds $\neg P \wedge Q$, which clashes with the premise, so $\neg Q$, and $\vee$I. That derivation has 11 lines, and its classical step is the LEM line.

Marking scheme. (10 points)

- 2: step 2 postpones the disjunctive goal, with the reason (neither side is in hand).
- 2: steps 3 and 4 find that the premise can be used only by producing $\neg P \wedge Q$ for a clash.
- 2: step 5 is chosen: RAA on the whole goal, or LEM on P. (LEM on Q also works; our route with it then uses a second classical step, an RAA for P in the case Q; award it.) A classical derived rule used at step 3, such as DM (d) on the premise followed by $\neg\neg$E, is also accepted; its DM (d) and $\neg\neg$E lines are the classical lines, and naming either earns the point.
- 3: the inner boxes and their clashes are correct, and every citation is available.
- 1: the classical line is named.

Fatal errors (0 for the step): $\vee$I as the last line, from a disjunct derived from the premise alone; on the RAA route, an RAA box on the whole goal opened with anything but $\neg(P \vee \neg Q)$.

Exercise. C4: Two sequents that look alike

Derive each sequent, and say which of your two derivations needed the classical rule.

(a) $\neg(P \wedge Q),\ P \vdash \neg Q$.

(b) $\neg(P \wedge \neg Q),\ P \vdash Q$.

Then explain why, in (b), the natural ¬I box on $\neg Q$ ends at $\neg\neg Q$ and not at Q.

Solution.

(a) The goal is a negation: a ¬I box on Q, in which $P \wedge Q$ clashes with the first premise.

1	$\neg(P \wedge Q)$	
2	P	
3	$\quad Q$	asm, for ¬I
4	$\quad P \wedge Q$	∧I 2, 3
5	$\quad \bot$	¬E 4, 1
6	$\neg Q$	¬I 3–5

(b) The goal Q is a letter, and the resources give nothing directly, so step 5: RAA, a box on $\neg Q$, in which $P \wedge \neg Q$ clashes with the first premise.

1	$\neg(P \wedge \neg Q)$	
2	P	
3	$\quad \neg Q$	asm, for RAA
4	$\quad P \wedge \neg Q$	∧I 2, 3
5	$\quad \bot$	¬E 4, 1
6	Q	RAA 3–5

The derivation of (b) needed the classical rule, at line 6; the derivation of (a) did not. The two have the same shape, and the difference is the goal. In (a) it is a negation, which ¬I builds. In (b) it is a letter, and the box that reaches $\bot$ opens with $\neg Q$; ¬I writes the negation of a box's assumption, so closing that box by ¬I gives $\neg\neg Q$, and our derivation reaches Q with the classical rule instead: RAA in place of ¬I (or ¬¬E after it).

3.2 Four Kinds of Task

The method of Section 3.1 answers one kind of question: here are premises, here is a conclusion, find a derivation. Mathematics asks others as well:

- prove a statement from nothing;
- show that two statements say the same thing;
- show that some conditions cannot all be met at once;
- decide whether a conclusion follows at all, when nobody says whether it does.

Each of these changes where a search starts and what counts as its end. This section takes the four tasks in turn. Underneath, it finds, they are one question.

3.2.1 Four tasks, one method

Let the task choose the search. The four tasks use the same rules and the same method. What differs is the first move and the finish line.

Task	The claim	Where the search starts	What ends it
theorem	$\vdash \varphi$	step 2 on the goal; every resource is an assumption the goal asks for	the goal, reached outside every box
equivalence	$\varphi \dashv\vdash \psi$ (Definition 2.4.2)	two searches, one from each side	two derivations
inconsistency	$\Gamma \vdash \bot$	step 4: the goal is $\bot$; look for a clash partner already on a line	$\bot$, outside every box
derive or break	$\Gamma \vdash \varphi$? (verdict withheld)	try to break it first	a derivation, or statements that make the counterexample set true

Only the last task leaves open which way it ends. The other three state their claim: a failed search leaves the question open, as Section 3.1's failed search did, and a break ends the task against the claim.

Quick check. Which task is each? (i) $\vdash P \vee \neg P$. (ii) $P \wedge Q \dashv\vdash Q \wedge P$. (iii) Show that P, $\neg P \vee Q$ and $\neg Q$ cannot all be accepted together. (iv) Does $P \to Q$, $Q \vdash P$ hold?

Solution.

(i) A theorem. (ii) An equivalence. (iii) An inconsistency: the claim P, $\neg P \vee Q$, $\neg Q \vdash \bot$. (iv) Derive or break; it breaks, since it is affirming the consequent (Section 1.3).

3.2.2 Theorems

A **theorem** of logic is a statement derivable from no premises, $\vdash \varphi$ (Definition 1.2.3 with the list empty).

Example. A statement that implies a contradiction

Find a derivation of $\vdash (P \to (Q \wedge \neg Q)) \to \neg P$.

In words: a statement that implies a contradiction is false. Every proof that begins "suppose P" and ends "which is impossible" has this shape.

Solution. There are no premises, so step 1 writes only the goal. Step 2 opens a box on $P \to (Q \wedge \neg Q)$ with the goal $\neg P$, and then, inside it, a box on P with the goal $\bot$:

1	$P \to (Q \wedge \neg Q)$	asm, for $\to$I
	⋮	
7	$\neg P$	
8	$(P \to (Q \wedge \neg Q)) \to \neg P$	$\to$I 1–7

1	$P \to (Q \wedge \neg Q)$	asm, for $\to$I
2	P	asm, for $\neg$I
	⋮	
6	$\bot$	
7	$\neg P$	$\neg$I 2–6
8	$(P \to (Q \wedge \neg Q)) \to \neg P$	$\to$I 1–7

Now there are two resources, both assumptions the goal asked for. Step 3 fires the first on the second, giving $Q \wedge \neg Q$, and splits it by $\wedge$E. Step 4 finds the clash.

1	$P \to (Q \wedge \neg Q)$	asm, for $\to$I
2	P	asm, for $\neg$I
3	$Q \wedge \neg Q$	$\to$E 1, 2
4	Q	$\wedge$E 3
5	$\neg Q$	$\wedge$E 3
6	$\bot$	$\neg$E 4, 5
7	$\neg P$	$\neg$I 2–6
8	$(P \to (Q \wedge \neg Q)) \to \neg P$	$\to$I 1–7

A theorem has no premises, so the goal does all the early work. Once proved, it is a derived rule with nothing above the line (Definition 2.3.1): any derivation, from any premises, may write an instance of it as a line, as derivations already write LEM.

3.2.3 Equivalences

Example. Proving a disjunction by refuting one side

Show that $P \to (Q \vee R) \dashv\vdash (P \wedge \neg Q) \to R$.

This is how mathematics proves a disjunction: to show that $ab = 0$ implies $a = 0$ or $b = 0$, assume $ab = 0$ and $a \neq 0$, and show $b = 0$ (divide by a). The equivalence says that this changes nothing.

Solution. ($\Rightarrow$) The goal is an implication, so a box on $P \wedge \neg Q$. $\wedge$E splits it, the premise fires on P, and DS removes the refuted side.

1	$P \to (Q \vee R)$	
2	$\quad P \wedge \neg Q$	asm, for $\to$I
	$\quad \vdots$	
6	$\quad R$	
7	$(P \wedge \neg Q) \to R$	$\to$I 2–6

1	$P \to (Q \vee R)$	
2	$\quad P \wedge \neg Q$	asm, for $\to$I
3	$\quad P$	$\wedge$E 2
4	$\quad \neg Q$	$\wedge$E 2
5	$\quad Q \vee R$	$\to$E 1, 3
	$\quad \vdots$	
6	$\quad R$	
7	$(P \wedge \neg Q) \to R$	$\to$I 2–6

1	$P \to (Q \vee R)$	
2	$\quad P \wedge \neg Q$	asm, for $\to$I
3	$\quad P$	$\wedge$E 2
4	$\quad \neg Q$	$\wedge$E 2
5	$\quad Q \vee R$	$\to$E 1, 3
6	$\quad R$	DS 5, 4
7	$(P \wedge \neg Q) \to R$	$\to$I 2–6

($\Leftarrow$) The goal is an implication: a box on P, with the goal $Q \vee R$. That goal is a disjunction, so it is postponed (step 2). The premise waits for $P \wedge \neg Q$: P is in hand, $\neg Q$ is not, and nothing clashes. That is step 5, and the well-chosen statement is Q, whose negation the premise waits for: LEM on Q.

1	$(P \wedge \neg Q) \to R$	
2	$\quad P$	asm, for $\to$I
	$\quad \vdots$	
10	$\quad Q \vee R$	
11	$P \to (Q \vee R)$	$\to$I 2–10

1	$(P \wedge \neg Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad Q \vee \neg Q$	LEM
4	$\quad\quad Q$	asm, for $\vee$E
	$\quad\quad \vdots$	
5	$\quad\quad Q \vee R$	
6	$\quad\quad \neg Q$	asm, for $\vee$E
	$\quad\quad \vdots$	
9	$\quad\quad Q \vee R$	
10	$\quad Q \vee R$	$\vee$E 3, 4–5, 6–9
11	$P \to (Q \vee R)$	$\to$I 2–10

In the case Q, $\vee$I gives the left side. In the case $\neg Q$, $\wedge$I builds $P \wedge \neg Q$, the premise gives R, and $\vee$I gives the right side.

1	$(P \wedge \neg Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad Q \vee \neg Q$	LEM
4	$\quad\quad Q$	asm, for $\vee$E
5	$\quad\quad Q \vee R$	$\vee$I 4
6	$\quad\quad \neg Q$	asm, for $\vee$E
7	$\quad\quad P \wedge \neg Q$	$\wedge$I 2, 6
8	$\quad\quad R$	$\to$E 1, 7
9	$\quad\quad Q \vee R$	$\vee$I 8
10	$\quad Q \vee R$	$\vee$E 3, 4–5, 6–9
11	$P \to (Q \vee R)$	$\to$I 2–10

An equivalence is two searches, and they need not be alike. The direction ($\Rightarrow$) used a derived rule and no classical step. The direction ($\Leftarrow$) reached step 5, because its goal was a disjunction with no side in hand. When one direction is easy, do it first: it checks that the statements are worth comparing, and its lines often suggest the other direction's split.

3.2.4 Inconsistent lists

Sometimes the task is not to reach a conclusion but to show that some statements cannot be accepted together: a puzzle's conditions that no number meets, the hypotheses of a lemma that contradict each other, or the assumption of a proof by contradiction together with what is given. Section 2.5 (Exercise (B5: Derive or break, used later)) already read an RAA derivation as a derivation of $\bot$ from the counterexample set. A derivation of $\bot$ does this job, and the list it

starts from gets a name.

Definition 3.2.1. Inconsistent and Consistent Lists

Let Γ be a list of statements $\varphi_1, \ldots, \varphi_n$, where $n \geq 0$. The list Γ is **inconsistent** if $\Gamma \vdash \bot$: some derivation whose premises all belong to Γ has $\bot$ as its last line. It is **consistent** if it is not inconsistent, that is, if **no** derivation with premises in Γ ends with $\bot$.

In words: a list is inconsistent when a derivation from it reaches $\bot$ outside every box. The definition speaks only of derivations, and says nothing about whether the statements are true. The list $P, \neg P$ is inconsistent, by one $\neg$E line, and so is any list with $\bot$ as a member. To show a list inconsistent, exhibit a derivation. To show one consistent, one would have to rule out every derivation, and nothing in this chapter can do that.

Proposition 3.2.1. An Inconsistent List Proves Everything

For a list of statements Γ, the following are equivalent:

(i) Γ is inconsistent;
(ii) $\Gamma \vdash \psi$ for **every** statement ψ;
(iii) $\Gamma \vdash \chi$ and $\Gamma \vdash \neg\chi$ for **some** statement χ.

In words: an inconsistent list proves every statement, and a list that proves some statement and its negation is inconsistent. This is Proposition 2.1.1 grown from two premises to any inconsistent list. Its proof is in the fold of the next subsection.

Quick check. Without writing a derivation, explain why $P,\ \neg P \vdash Q \wedge \neg Q$.

Solution. The list $P, \neg P$ is inconsistent (one $\neg$E line), so by Proposition 3.2.1, (i) $\Rightarrow$ (ii), it derives every statement, $Q \wedge \neg Q$ included.

Example. Four conditions that no integer meets

Let n be an integer. A puzzle asks for an n such that:

(1) if n is even, then n is a multiple of 4;
(2) n is even or a multiple of 3;
(3) if n is a multiple of 3, then n is even;
(4) n is not a multiple of 4.

With E: "n is even", F: "n is a multiple of 4" and T: "n is a multiple of 3", show that the list $E \to F,\ E \vee T,\ T \to E,\ \neg F$ is inconsistent.

Solution. The goal is $\bot$. Look first for a clash partner already on a line (Section 3.1): line 4 is $\neg F$, so aim at F.

F is the conclusion of line 1, so the new goal is E, a letter. Line 3 gives E from T, but T is not in hand, and $E \vee T$ does not say which side holds. So use line 2 by $\vee$E, with both cases ending in E. The case E is a one-line box; the case T fires line 3.

1	$E \to F$	
2	$E \vee T$	
3	$T \to E$	
4	$\neg F$	
	⋮	
9	F	
10	$\bot$	$\neg$E 9, 4

1	$E \to F$	
2	$E \vee T$	
3	$T \to E$	
4	$\neg F$	
5	$\quad E$	asm, for $\vee$E
6	$\quad T$	asm, for $\vee$E
	$\quad$ ⋮	
7	$\quad E$	
8	E	$\vee$E 2, 5–5, 6–7
9	F	$\to$E 1, 8
10	$\bot$	$\neg$E 9, 4

1	$E \to F$	
2	$E \vee T$	
3	$T \to E$	
4	$\neg F$	
5	$\quad E$	asm, for $\vee$E
6	$\quad T$	asm, for $\vee$E
7	$\quad E$	$\to$E 3, 6
8	E	$\vee$E 2, 5–5, 6–7
9	F	$\to$E 1, 8
10	$\bot$	$\neg$E 9, 4

Read as a proof about n, the derivation shows that no integer meets all four conditions: by (2) and (3), n is even in either case; by (1) it is a multiple of 4; and (4) says it is not. Yet any three of the conditions can be met together: $n = 6$ meets all but (1), $n = 1$ all but (2), $n = 3$ all but (3), and $n = 12$ all but (4). The inconsistency belongs to the four together. The derivation does not say which condition to drop (the fold "Logic permits; it does not choose" in Section 3.1).

Quick check. Which integers between 1 and 10 meet conditions (1), (2) and (4) of Example (Four conditions that no integer meets), and which condition does each break?

Solution. 3 and 9: they are odd, so (1) promises nothing; they are multiples of 3, so (2) holds; and they are not multiples of 4. Each breaks (3).

Here are three more inconsistent lists, each with compound members, and one list that only looks inconsistent.

Example. Three inconsistent lists, and a look-alike

The list	The clash it reaches	The derivation of $\bot$, in brief
$(P \vee Q) \to R,\ P \wedge \neg R$	R, $\neg R$	P by $\wedge$E; $P \vee Q$ by $\vee$I; R by $\to$E; $\neg R$ by $\wedge$E; $\bot$ by $\neg$E (7 lines)
$P \leftrightarrow \neg P$	P, $\neg P$	a box on P: $\neg P$ by $\leftrightarrow$E, $\bot$ by $\neg$E; $\neg P$ by $\neg$I; then P by $\leftrightarrow$E, and $\neg P$ by $\leftrightarrow$E again from that P; $\bot$ by $\neg$E (8 lines)
$\neg(P \to Q),\ \neg P$	$P \to Q$, $\neg(P \to Q)$	$P \to Q$ by VI from $\neg P$ (Proposition 2.1.2); $\bot$ by $\neg$E (4 lines)

Not an instance. $P \to Q,\ P \to \neg Q$. The two members look opposed. But with P false (say P: "$0 = 1$") both are true, whatever Q says, so the claim $P \to Q,\ P \to \neg Q \vdash \bot$ breaks. What the list does derive is $\neg P$: from P it would give Q and $\neg Q$. This chapter does not call the list inconsistent, and it does not call it consistent either. Whether a break rules out every derivation of $\bot$ is the question the chapter ends on (Section 3.3).

The second row's derivation is written without a detour: the line $\neg P$ built by $\neg$I is cited by $\leftrightarrow$E, never taken apart at once by $\neg$E, so it stays, as Section 3.1's detour paragraph allows ("an introduced line cited elsewhere stays").

Warning (Not allowed: claims about inconsistency). Each claim below is wrong; the tag names what it misreads.

- "The list is inconsistent because one of its members is false." *Truth for derivation*: inconsistency is a derivation of $\bot$ (Definition 3.2.1), not a false member. The one-member list "$2 > 3$" has a false member, and no derivation of $\bot$ is offered.
- "$P \to Q$ and $P \to \neg Q$ contradict each other, so the list is inconsistent." *Looks for derivation*: nothing has been derived. Both members are true when P is false, and the claim of inconsistency breaks.
- "Line 7, inside the box of the assumption P, is $\bot$; so the premises are inconsistent." *A boxed line*: that $\bot$ depends on P. It shows the list Γ, P inconsistent, and so $\Gamma \vdash \neg P$. The last line must lie outside every box (Definition 1.4.2).
- "A search for $\bot$ stalled, so the list is consistent." *A stall for a proof*: a stalled search proves nothing (Section 3.1). Consistency is the absence of **every** derivation of $\bot$.
- "The list is inconsistent, so its first member is false." *The list for a member*: inconsistency belongs to the list. It says that the members cannot all be accepted together,

not which one fails (compare the four conditions above).

3.2.5 The counterexample set, derived

Proving and refuting meet in one list. The counterexample set of a claim $\Gamma \vdash \varphi$ is the list $\Gamma, \neg\varphi$ (Definition 1.1.3). Breaking the claim makes that list true. Deriving the claim, it turns out, is showing that list inconsistent.

Proposition 3.2.2. Derivable Exactly When the Counterexample Set Is Inconsistent

Let Γ be a list of statements and φ a statement.

(a) $\Gamma \vdash \neg\varphi$ if and only if the list Γ, φ is inconsistent.

(b) $\Gamma \vdash \varphi$ if and only if the list $\Gamma, \neg\varphi$ is inconsistent.

In words: (b) says that a claim is derivable exactly when its counterexample set is inconsistent, and (a) says the same for a negated conclusion. In the proofs, the box of (a) closes by ¬I and that of (b) by RAA: Section 2.5's distinction between the two rules, stated about whole derivations. Section 1.1's test of a form, try to break it, now has a partner on the same list: make the counterexample set true, or derive $\bot$ from it.

With Γ empty and φ the statement $P \leftrightarrow \neg P$, part (a) says that the second row of Example (Three inconsistent lists, and a look-alike) and Lemma 2.4.2, $\vdash \neg(P \leftrightarrow \neg P)$, are the same fact.

Second reading · *The proofs: append a line, or copy a derivation into a box*

Proof. Proof of Proposition 3.2.1. (i) ⇒ (ii): a derivation of $\bot$ with premises in Γ, followed by the line ψ by $\bot$E, is a derivation of ψ. Its last line lies in no box. (ii) ⇒ (iii): apply (ii) to any χ and to $\neg\chi$. (iii) ⇒ (i): the sequent $\chi, \neg\chi \vdash \bot$ has a one-step derivation (¬E), so by Lemma 2.3.3 (c), $\Gamma \vdash \bot$.

■

Proof. Proof of Proposition 3.2.2. (b), (⇒): from $\Gamma \vdash \varphi$, Lemma 2.3.3 (a) gives $\Gamma, \neg\varphi \vdash \varphi$. The one-line derivation of the premise $\neg\varphi$ gives $\Gamma, \neg\varphi \vdash \neg\varphi$. By Lemma 2.3.3 (c) with $\varphi, \neg\varphi \vdash \bot$, we get $\Gamma, \neg\varphi \vdash \bot$.

(b), (⇐): let D be a derivation of $\bot$ with premises in $\Gamma, \neg\varphi$. If the last line of D is a premise in Γ, then $\bot$ is a member of Γ, and $\bot$E gives φ. (The last line cannot be the premise $\neg\varphi$, since $\neg\varphi$ is not $\bot$.) Otherwise, write as premises those premises of D that belong to Γ. Open a box with the assumption $\neg\varphi$. Inside it, copy the lines of D that are not premises, in order, with their boxes, renumbering citations; a citation of the premise $\neg\varphi$ now cites the assumption. Close the box after the copy of the last line of D, and write φ by RAA.

Every copied citation is legal. A cited premise of Γ lies above the box, in no box. The assumption is available throughout the box. The boxes of D sit inside the new box as they sat in D, so every line and every box that was available at a citing line stays available (the scope rule of Definition 1.4.2). The last line of D lies in no box of D, so its copy is a line of

the new box itself, as RAA requires. If D never cites $\neg\varphi$, nothing changes.

Part (a) is the same, with φ in place of $\neg\varphi$ and $\neg$I in place of RAA. If the last line of D is the premise φ itself (possible only when φ is $\bot$), nothing is copied: the box is its one assumption line, and $\neg$I closes it at once. In ($\Rightarrow$), $\Gamma, \varphi \vdash \varphi$ and, by weakening, $\Gamma, \varphi \vdash \neg\varphi$; chaining with $\neg$E finishes.

■

The copying is the move of Section 1.4's Exercise (C4: The assumptions a line depends on, used later) (b), which turns open assumptions into premises, run backwards.

Quick check. By Proposition 3.2.2, which list must be shown inconsistent to prove $\vdash \neg(P \wedge \neg P)$? Show it.

Solution. By (a), with Γ empty and φ the statement $P \wedge \neg P$, the list is $P \wedge \neg P$:

1	$P \wedge \neg P$	
2	P	$\wedge$E 1
3	$\neg P$	$\wedge$E 1
4	$\bot$	$\neg$E 2, 3

3.2.6 One task in four costumes

Every task of this section is a claim that something is derivable, and by Proposition 3.2.2 every such claim is a claim that a list is inconsistent.

Task	As a claim of derivability	As a claim of inconsistency	It breaks when some choice of statements makes
theorem	$\vdash \varphi$	the list $\neg\varphi$	φ false
equivalence	$\varphi \vdash \psi$ and $\psi \vdash \varphi$	the lists $\varphi, \neg\psi$ and $\psi, \neg\varphi$	one of φ, ψ true and the other false
inconsistency	$\Gamma \vdash \bot$	the list Γ	every member of Γ true
derive or break	$\Gamma \vdash \varphi$	the list $\Gamma, \neg\varphi$	every member of $\Gamma, \neg\varphi$ true

An equivalence is also a theorem: $\varphi \dashv\vdash \psi$ holds exactly when $\vdash \varphi \leftrightarrow \psi$. Copy each of the two derivations into a box, as in the fold above, and close both boxes by $\leftrightarrow$I. Conversely, add the premise φ to a derivation of $\varphi \leftrightarrow \psi$ and write ψ by $\leftrightarrow$E, and likewise for ψ.

So the four tasks are one: is this list inconsistent? Each costume starts the search in a different

place, which is why the task chooses the search. And each row has a second half, the last column: what to do when the list might instead be made true. That is the fourth task.

3.2.7 Derive or break

Section 1.5 listed claims without a verdict ("Derive it or break it"). Here the format becomes a task with its own first move: **decide before you search**.

Try to break it first. Spend a few minutes trying to make the counterexample set true, starting where the conclusion is false (Section 1.1, "Where to look first"). Either some choice of statements works, and the claim breaks; or every attempt is forced into a clash. Then search for a derivation. The forced attempt usually suggests a plan for it: the statements the attempt had to make true suggest the assumptions of the derivation's boxes, and the clash it met suggests the derivation's $\bot$.

Example. Two claims, one connective apart

Derive or break each claim.

(a) $P \to (Q \to R),\ P \to \neg R \vdash P \to \neg Q$.
(b) $P \to (Q \vee R),\ P \to \neg R \vdash P \to \neg Q$.

Solution. (a), trying to break it. The conclusion is false only with P true and $\neg Q$ false, that is, Q true. Then the first premise needs $Q \to R$ true, so R true. The second needs $\neg R$ true, so R false. Every choice was forced, and the forcing ends in a clash. So we search.

(a), the derivation. The attempt's plan: a box on P (the conclusion's hypothesis), a box on Q (the statement whose negation is wanted), and the clash R, $\neg R$. Before the box on Q opens, both premises fire on P: lines that need nothing new (Section 3.1).

1	$P \to (Q \to R)$	
2	$P \to \neg R$	
3	$\quad P$	asm, for $\to$I
	$\quad \vdots$	
9	$\quad \neg Q$	
10	$P \to \neg Q$	$\to$I 3–9

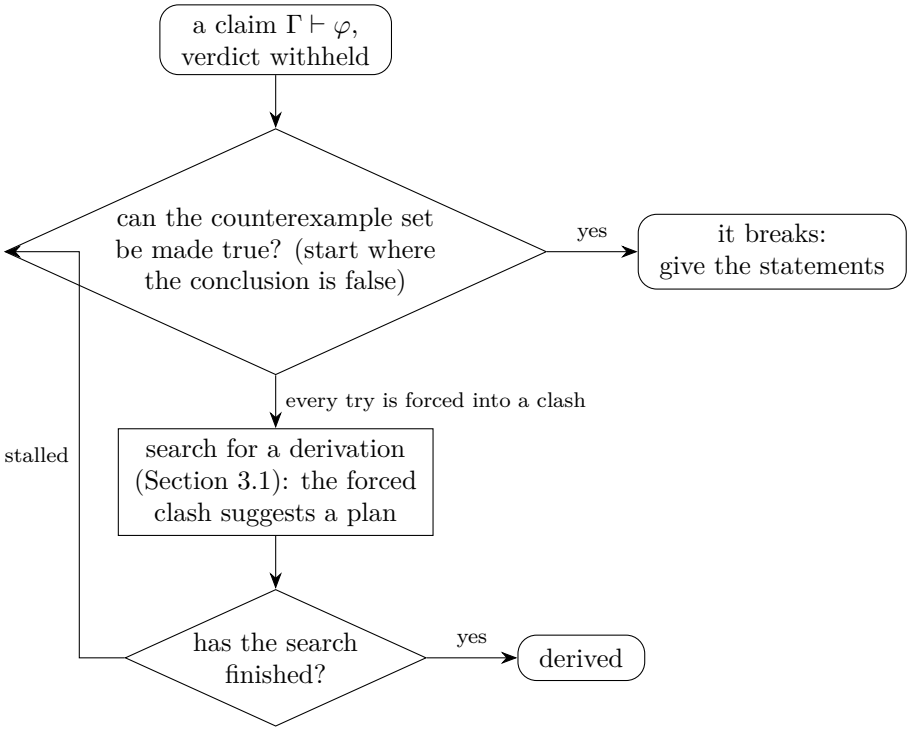
1	$P \to (Q \to R)$	
2	$P \to \neg R$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q \to R$	$\to$E 1, 3
5	$\quad \neg R$	$\to$E 2, 3
6	$\quad\quad Q$	asm, for $\neg$I
	$\quad\quad \vdots$	
8	$\quad\quad \bot$	
9	$\quad \neg Q$	$\neg$I 6–8
10	$P \to \neg Q$	$\to$I 3–9

1	$P \to (Q \to R)$	
2	$P \to \neg R$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q \to R$	$\to$E 1, 3
5	$\quad \neg R$	$\to$E 2, 3
6	$\quad\quad Q$	asm, for $\neg$I
7	$\quad\quad R$	$\to$E 4, 6
8	$\quad\quad \bot$	$\neg$E 7, 5
9	$\quad \neg Q$	$\neg$I 6–8
10	$P \to \neg Q$	$\to$I 3–9

(b), trying to break it. The same start: P true, Q true. Now the first premise asks only that $Q \vee R$ be true, and Q already makes it so. R is free, and the second premise asks for R false. Nothing clashes: with P and Q true and R false, every member of the counterexample set $P \to (Q \vee R),\ P \to \neg R,\ \neg(P \to \neg Q)$ is true. With statements about a real number x, take P: "$x > 0$", Q: "$x > 1$", R: "$x > 3$", at $x = 2$. The claim breaks.

The two claims differ in one connective. Trying to break each found its verdict, and for (a) it also found the derivation. The method of Section 3.1 would have found (a) without the attempt. On (b) the method of Section 3.1 could only search, and a stall is not a verdict; the break is.

The task as a chart. The loop at the bottom is honest: nothing in this chapter promises that it ends.



Quick check. Change one symbol of the first premise of claim (b) in Example (Two claims, one connective apart) so that the claim can be derived, and derive it.

Solution. For example, replace $\vee$ by $\to$: that is claim (a), derived above. Or replace $\vee$ by $\wedge$: from P, the premises give $Q \wedge R$ and $\neg R$, which clash.

1	$P \to (Q \wedge R)$	
2	$P \to \neg R$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q \wedge R$	$\to$E 1, 3
5	$\quad R$	$\wedge$E 4
6	$\quad \neg R$	$\to$E 2, 3
7	$\quad \bot$	$\neg$E 5, 6
8	$\quad \neg Q$	$\bot$E 7
9	$P \to \neg Q$	$\to$I 3–8

Replacing $\vee$ by $\leftrightarrow$, or Q by R, works too. For $\leftrightarrow$: from P, the premises give $Q \leftrightarrow R$ and $\neg R$; a box on Q gives R by $\leftrightarrow$E, which clashes with $\neg R$, so $\neg Q$ by $\neg$I. For Q replaced by R: from P, $R \vee R$ gives R by $\vee$E, which clashes with $\neg R$, so $\neg Q$ by $\bot$E.

Remark (Looking ahead). Chapter 4 calls a list whose members can all be made true at once *satisfiable*, and proves that the satisfiable lists are exactly the consistent ones.

3.2.8 Exercises

3.2.8.1 A. Check your understanding

Exercise. A1

(a) When is a list of statements inconsistent?
(b) By Proposition 3.2.2, which list is inconsistent exactly when $P \to Q,\ Q \to R \vdash P \to R$?
(c) True or false: every list that has $\bot$ as a member is inconsistent.
(d) True or false: if Γ is inconsistent, so is every list Γ, ψ.
(e) Name the task and the first move: "show that $P \wedge Q \dashv\vdash Q \wedge P$".

Solution.

(a) When $\bot$ is derivable from it (Definition 3.2.1).
(b) The counterexample set $P \to Q,\ Q \to R,\ \neg(P \to R)$.
(c) True: the one-line derivation whose only line is the premise $\bot$ ends with $\bot$.

(d) True, by Lemma 2.3.3 (a): the derivation of $\bot$ from Γ has its premises in Γ, ψ.

(e) An equivalence: two derivations, each starting with $\wedge$E on its premise. Each is Proposition 1.2.1, with the letters exchanged for the second.

Exercise. A2: Drill

The parts run from naming the task, through one step and a missing piece, to claims to judge and short derivations.

(a) Which task is "show $\vdash (P \wedge Q) \to P$"?

(b) "Show that $P \to \neg Q$ and $Q \to \neg P$ are provably equivalent." Which two derivations does this ask for?

(c) Write "P, $P \to Q$ and $\neg Q$ cannot all be accepted together" as a sequent.

(d) A search for $\bot$ from $P \wedge \neg Q$, $P \to Q$ begins. Write its first three lines after the premises.

(e) Which list must be shown inconsistent to prove $P \to Q \vdash \neg Q \to \neg P$ by Proposition 3.2.2 (b)?

(f) The list Q, $\neg P$, $\neg\neg R$ is the counterexample set of which claims?

(g) "$P \to Q$, $P \to \neg Q$ is inconsistent, since its members contradict each other." Is this correct?

(h) "Inside the box of the assumption P, line 6 is $\bot$; so the premises are inconsistent." Is this correct?

(i) "A search for $\bot$ from $P \vee Q$, $\neg P$ stalled; so the list is consistent." Is this correct?

(j) "Since $P, \neg P \vdash \bot$, the list $P, \neg P$ derives both Q and $\neg Q$." Is this correct?

(k) Derive $\bot$ from $P \to Q$, $Q \to R$, P, $\neg R$.

(l) Derive $\bot$ from $P \vee Q$, $\neg P$, $\neg Q$.

Solution.

(a) A theorem.

(b) $P \to \neg Q \vdash Q \to \neg P$ and $Q \to \neg P \vdash P \to \neg Q$.

(c) P, $P \to Q$, $\neg Q \vdash \bot$.

(d) 3 P, $\wedge$E 1; 4 Q, $\to$E 2, 3; 5 $\neg Q$, $\wedge$E 1. (Then 6 $\bot$, $\neg$E 4, 5.)

(e) $P \to Q$, $\neg(\neg Q \to \neg P)$.

(f) Of $Q, \neg P \vdash \neg R$, with $\neg\neg R$ the negated conclusion, and of $Q, \neg\neg R \vdash P$, with $\neg P$ the negated conclusion. Q is not a negation, so it is not the negated conclusion of any claim.

(g) No. Both members are true when P is false, and nothing has been derived.

(h) No. It shows the premises together with P inconsistent, so the premises derive $\neg P$.

(i) No: a stalled search proves nothing.

(j) Yes: Proposition 3.2.1, (i) $\Rightarrow$ (ii).

(k) 5 Q, $\to$E 1, 3; 6 R, $\to$E 2, 5; 7 $\bot$, $\neg$E 6, 4.

(l) 4 P, assumption; 5 $\bot$, $\neg$E 4, 2; 6 Q, assumption; 7 $\bot$, $\neg$E 6, 3; 8 $\bot$, $\vee$E 1, 4–5, 6–7.

3.2.8.2 B. Practice

Exercise. B1: Three theorems

Derive each theorem, and say which step of Section 3.1's method opens each box.

(a) $\vdash \neg(P \wedge \neg P) \wedge \neg\neg(P \vee \neg P)$.
(b) $\vdash ((P \to Q) \wedge (P \vee Q)) \to Q$.
(c) $\vdash ((P \to Q) \to Q) \to ((Q \to P) \to P)$.

Solution.

(a)

1	$P \wedge \neg P$	asm, for $\neg$I
2	P	$\wedge$E 1
3	$\neg P$	$\wedge$E 1
4	$\bot$	$\neg$E 2, 3
5	$\neg(P \wedge \neg P)$	$\neg$I 1–4
6	$\neg(P \vee \neg P)$	asm, for $\neg$I
7	P	asm, for $\neg$I
8	$P \vee \neg P$	$\vee$I 7
9	$\bot$	$\neg$E 8, 6
10	$\neg P$	$\neg$I 7–9
11	$P \vee \neg P$	$\vee$I 10
12	$\bot$	$\neg$E 11, 6
13	$\neg\neg(P \vee \neg P)$	$\neg$I 6–12
14	$\neg(P \wedge \neg P) \wedge \neg\neg(P \vee \neg P)$	$\wedge$I 5, 13

Boxes 1 and 6 come from step 2: each goal is a negation. Box 7 comes from step 2 inside step 4: step 4 chooses to manufacture $P \vee \neg P$ from its side $\neg P$, so that it clashes with line 6, and the goal $\neg P$, a negation, opens the box. The second conjunct needed no classical rule, although $P \vee \neg P$ itself does in this book's derivation (Proposition 2.5.3).

(b)

1	$(P \to Q) \wedge (P \vee Q)$	asm, for $\to$I
2	$P \to Q$	$\wedge$E 1
3	$P \vee Q$	$\wedge$E 1
4	P	asm, for $\vee$E
5	Q	$\to$E 2, 4
6	Q	asm, for $\vee$E
7	Q	$\vee$E 3, 4–5, 6–6
8	$((P \to Q) \wedge (P \vee Q)) \to Q$	$\to$I 1–7

Box 1 comes from step 2. Boxes 4 and 6 come from step 3: the resource $P \vee Q$ is split

by $\vee$E, both cases ending in the goal Q (the case Q is a one-line box). Lines 2 and 3 need nothing new. Read aloud: if P implies Q, and P or Q holds, then Q holds.

(c)

1	$(P \to Q) \to Q$	asm, for $\to$I
2	$\quad Q \to P$	asm, for $\to$I
3	$\quad\quad \neg P$	asm, for RAA
4	$\quad\quad P \to Q$	VI 3
5	$\quad\quad Q$	$\to$E 1, 4
6	$\quad\quad P$	$\to$E 2, 5
7	$\quad\quad \bot$	$\neg$E 6, 3
8	$\quad P$	RAA 3–7
9	$(Q \to P) \to P$	$\to$I 2–8
10	$((P \to Q) \to Q) \to ((Q \to P) \to P)$	$\to$I 1–9

Boxes 1 and 2 come from step 2. Box 3 comes from step 5: the goal P is a letter, and the resources wait for $P \to Q$, which nothing in hand gives.

Exercise. B2: Two directions

Show that $(P \wedge Q) \to R \dashv\vdash (P \wedge \neg R) \to \neg Q$.

(a) The search for ($\Leftarrow$) has reached the stage below. *Name the move, then complete it.*

(b) Derive ($\Rightarrow$).

(c) Which direction used a classical rule, and why did the other not need one?

1	$(P \wedge \neg R) \to \neg Q$	
2	$\quad P \wedge Q$	asm, for $\to$I
	$\quad \vdots$	
9	$\quad R$	
10	$(P \wedge Q) \to R$	$\to$I 2–9

Solution.

(a) First the lines that need nothing new, P and Q by $\wedge$E. Then the move is step 5: the current goal R is a letter, no resource gives it, and nothing clashes, so assume $\neg R$ for RAA.

	1	$(P \wedge \neg R) \to \neg Q$	
	2	$\quad P \wedge Q$	asm, for $\to$I
	3	$\quad P$	$\wedge$E 2
	4	$\quad Q$	$\wedge$E 2
	5	$\quad\quad \neg R$	asm, for RAA
	6	$\quad\quad P \wedge \neg R$	$\wedge$I 3, 5
	7	$\quad\quad \neg Q$	$\to$E 1, 6
	8	$\quad\quad \bot$	$\neg$E 4, 7
	9	$\quad R$	RAA 5–8
	10	$(P \wedge Q) \to R$	$\to$I 2–9

(b)

	1	$(P \wedge Q) \to R$	
	2	$\quad P \wedge \neg R$	asm, for $\to$I
	3	$\quad P$	$\wedge$E 2
	4	$\quad \neg R$	$\wedge$E 2
	5	$\quad\quad Q$	asm, for $\neg$I
	6	$\quad\quad P \wedge Q$	$\wedge$I 3, 5
	7	$\quad\quad R$	$\to$E 1, 6
	8	$\quad\quad \bot$	$\neg$E 7, 4
	9	$\quad \neg Q$	$\neg$I 5–8
	10	$(P \wedge \neg R) \to \neg Q$	$\to$I 2–9

(c) ($\Leftarrow$) used RAA at line 9, because its goal R is a letter. ($\Rightarrow$) needed none, because its goal $\neg Q$ is a negation, which $\neg$I builds. This is a claim about our two derivations only.

In words: to prove "if P and Q, then R", it is enough to prove "if P and not R, then not Q", a contrapositive in one hypothesis.

Exercise. B3: Derive or break, used later

For each claim, derive it or break it. Try to break it first. To break a claim, give statements that make its counterexample set true.

(a) $\neg P \vee \neg Q \vdash \neg(P \vee Q)$
(b) $\neg(P \leftrightarrow Q),\ P \vdash \neg Q$
(c) $P \to Q,\ \neg P \to (R \vee Q),\ \neg R \vdash Q$
(d) $P \to Q,\ \neg P \to (R \vee Q),\ \neg R \vdash P$
(e) $(P \to Q) \to R,\ \neg R \vdash P$
(f) $(P \to Q) \to R,\ \neg R \vdash Q$

(g) $P \vee Q,\ P \vdash \neg Q$
(h) $P \vee Q,\ \neg(P \wedge Q),\ P \vdash \neg Q$
(i) $\neg(P \vee Q) \to P \vdash P \vee Q$

Solution. The verdict comes first, then the derivation or the break. In every break, x is a real number.

(a) Breaks. Take P: "$x > 2$", Q: "$x > 0$", at $x = 1$. $\neg P$ is true, so $\neg P \vee \neg Q$ is true; Q is true, so $P \vee Q$ is true and $\neg(P \vee Q)$ is false. (The De Morgan law with this premise is $\neg P \vee \neg Q \vdash \neg(P \wedge Q)$, Proposition 2.5.5 (c); here the wrong connective sits inside the negation.)

(b) Derive.

1	$\neg(P \leftrightarrow Q)$	
2	P	
3	$\quad Q$	asm, for $\neg$I
4	$\quad\quad P$	asm, for $\leftrightarrow$I
5	$\quad\quad Q$	R 3
6	$\quad\quad Q$	asm, for $\leftrightarrow$I
7	$\quad\quad P$	R 2
8	$\quad P \leftrightarrow Q$	$\leftrightarrow$I 4–5, 6–7
9	$\quad \bot$	$\neg$E 8, 1
10	$\neg Q$	$\neg$I 3–9

This is the first box of Exercise (B5: Derive or break, used later) (f).

(c) Derive. Step 5: the goal Q is a letter, and the first two premises wait for P and for $\neg P$, so LEM on P. DS matches as stated in Section 2.3, refuting the **left** disjunct R.

1	$P \to Q$	
2	$\neg P \to (R \vee Q)$	
3	$\neg R$	
4	$P \vee \neg P$	LEM
5	$\quad P$	asm, for $\vee$E
6	$\quad Q$	$\to$E 1, 5
7	$\quad \neg P$	asm, for $\vee$E
8	$\quad R \vee Q$	$\to$E 2, 7
9	$\quad Q$	DS 8, 3
10	Q	$\vee$E 4, 5–6, 7–9

(d) Breaks. Take P: "$x > 5$", Q: "$x > 0$", R: "$x > 10$", at $x = 1$. P is false, so $P \to Q$ is true; $\neg P$ and Q are true, so $\neg P \to (R \vee Q)$ is true; $\neg R$ is true; and P is false. Trying to break it first: with P false, the premises force Q true and R false, and nothing

clashes. The tempting step from "P gives Q" and "Q holds", which (c) derives, to P is affirming the consequent (Section 1.3).

(e) Derive.

1	$(P \to Q) \to R$	
2	$\neg R$	
3	$\neg(P \to Q)$	MT 1, 2
4	$\quad \neg P$	asm, for RAA
5	$\quad P \to Q$	VI 4
6	$\quad \bot$	¬E 5, 3
7	P	RAA 4–6

(f) Breaks. Take P: "$x > 0$", Q: "$x > 2$", R: "$x > 3$", at $x = 1$. $P \to Q$ is false, so $(P \to Q) \to R$ is true; $\neg R$ is true; and Q is false.

(g) Breaks; it is affirming a disjunct (Section 1.6). Take P: "$x > 1$", Q: "$x > 2$", at $x = 3$: $P \vee Q$ and P are true, and $\neg Q$ is false.

(h) Derive. The premise $P \vee Q$ is not needed.

1	$P \vee Q$	
2	$\neg(P \wedge Q)$	
3	P	
4	$\quad Q$	asm, for ¬I
5	$\quad P \wedge Q$	∧I 3, 4
6	$\quad \bot$	¬E 5, 2
7	$\neg Q$	¬I 4–6

(i) Derive. Step 2 postpones the disjunction. The premise waits for $\neg(P \vee Q)$, which RAA on the goal supplies; step 4 then manufactures the clash partner $P \vee Q$ from line 3 by ∨I, a manufactured clash and not a detour.

1	$\neg(P \vee Q) \to P$	
2	$\quad \neg(P \vee Q)$	asm, for RAA
3	$\quad P$	→E 1, 2
4	$\quad P \vee Q$	∨I 3
5	$\quad \bot$	¬E 4, 2
6	$P \vee Q$	RAA 2–5

Read aloud: a statement whose failure would force one of its own sides holds.

(c) and (d), and (e) and (f), are look-alike pairs: each pair has the same premises and differs in the conclusion. (h) is (g) repaired by "not both". (a) is a De Morgan look-alike with the wrong connective inside.

Exercise. B4: Derive $\bot$, or make every member true, used later

For each list, either derive $\bot$ from it, or give statements that make every member true at once.

(a) $P \to Q,\ Q \to \neg P,\ P$
(b) $P \lor Q,\ P \to \neg Q,\ Q \to \neg P$
(c) $\neg(P \to Q),\ Q \lor \neg P$
(d) $P \to (Q \land \neg Q)$
(e) $\neg(P \lor Q),\ P \lor (Q \land R)$
(f) $\neg(P \land Q),\ \neg(P \land \neg Q),\ \neg(\neg P \land Q)$

Solution.

(a) Derive $\bot$: 4 Q, $\to$E 1, 3; 5 $\neg P$, $\to$E 2, 4; 6 $\bot$, $\neg$E 3, 5.

(b) True at once. Take P: "$x > 0$", Q: "$x < 0$", at $x = 1$: $P \lor Q$ is true; $\neg Q$ is true, so $P \to \neg Q$ is true; and $Q \to \neg P$ is true, since Q is false.

(c) Derive $\bot$.

1	$\neg(P \to Q)$	
2	$Q \lor \neg P$	
3	$\quad Q$	asm, for $\lor$E
4	$\quad\quad P$	asm, for $\to$I
5	$\quad\quad Q$	R 3
6	$\quad P \to Q$	$\to$I 4–5
7	$\quad \bot$	$\neg$E 6, 1
8	$\quad \neg P$	asm, for $\lor$E
9	$\quad P \to Q$	VI 8
10	$\quad \bot$	$\neg$E 9, 1
11	$\bot$	$\lor$E 2, 3–7, 8–10

(d) True at once, with P false (P: "$0 = 1$", any Q). A member that mentions a contradiction is not itself a clash: with P false every member is true, so the claim that the list derives $\bot$ breaks.

(e) Derive $\bot$.

1	$\neg(P \vee Q)$	
2	$P \vee (Q \wedge R)$	
3	$\quad P$	asm, for $\vee$E
4	$\quad P \vee Q$	$\vee$I 3
5	$\quad \bot$	$\neg$E 4, 1
6	$\quad Q \wedge R$	asm, for $\vee$E
7	$\quad Q$	$\wedge$E 6
8	$\quad P \vee Q$	$\vee$I 7
9	$\quad \bot$	$\neg$E 8, 1
10	$\bot$	$\vee$E 2, 3–5, 6–9

(f) True at once, with P and Q both false (P: "$1 > 2$", Q: "$2 > 3$"): each conjunction has a false member, so each negation is true. Adding a fourth member $\neg(\neg P \wedge \neg Q)$ gives a list from which $\bot$ can be derived: the four exclude every way for P and Q to be. Use LEM on P and, inside each case, on Q; each of the four cases builds the conjunction that one member negates, by $\wedge$I, and clashes with it.

Section 4.7 shows that a list made true at once, as in (b), (d) and (f), is consistent.

3.2.8.3 C. Going deeper

Exercise. C1: Exit check: the counterexample set, used later

Let Γ be the list $P \to (Q \vee R)$, $Q \to S$, $R \to S$.

(a) Write the counterexample set of $\Gamma \vdash P \to S$.

(b) Show that it is inconsistent.

(c) What does Proposition 3.2.2 conclude? Give also a direct derivation of $\Gamma \vdash P \to S$, and compare the two.

(d) Replace one premise so that the claim breaks, and break it.

Solution.

(a) $P \to (Q \vee R)$, $Q \to S$, $R \to S$, $\neg(P \to S)$.

(b) The goal is $\bot$, and the clash partner $\neg(P \to S)$ is already on a line (Section 3.1), so aim at $P \to S$.

1	$P \to (Q \vee R)$	
2	$Q \to S$	
3	$R \to S$	
4	$\neg(P \to S)$	
5	$\quad P$	asm, for $\to$I
6	$\quad Q \vee R$	$\to$E 1, 5
7	$\quad\quad Q$	asm, for $\vee$E
8	$\quad\quad S$	$\to$E 2, 7
9	$\quad\quad R$	asm, for $\vee$E
10	$\quad\quad S$	$\to$E 3, 9
11	$\quad S$	$\vee$E 6, 7–8, 9–10
12	$P \to S$	$\to$I 5–11
13	$\bot$	$\neg$E 12, 4

(c) By part (b) of the proposition, $\Gamma \vdash P \to S$. The direct derivation is lines 1–3 and 5–12 of (b), renumbered: 11 lines. The route through the counterexample set adds the premise $\neg(P \to S)$ and one $\neg$E line, and gains nothing here. It pays when the goal gives step 2 nothing (a letter, or a disjunction with no side in hand), and then it is step 5.

(d) For instance, replace $R \to S$ by $R \to T$. Take P: "$x > 0$", Q: "$x > 5$", R: "$x < 2$", S: "$x > 3$", T: "$x < 10$", at $x = 1$. P and R are true, so $Q \vee R$ and $P \to (Q \vee R)$ are true; $Q \to S$ is true, since Q is false; $R \to T$ is true, since T is true; and $P \to S$ is false, since S is false. So every member of the new counterexample set is true. Any replacement that a choice of statements breaks is accepted.

Marking scheme. (10 points)

- 2: the counterexample set, with $\neg(P \to S)$, not $\neg P \to S$ or $P \to \neg S$.
- 3: the derivation of $\bot$: $P \to S$ derived (by a box whose cases of $Q \vee R$ both end in S, or by Example (Cases on a hypothesis are two implications) and Proposition 1.4.1), and $\neg$E outside every box.
- 2: the conclusion drawn by the proposition, and the direct derivation.
- 3: (d), a changed premise, with statements checked against all four members of the new counterexample set.

Fatal errors (0 for the step): $\bot$ written inside the box of P and taken as the end; a "break" in (d) that leaves a premise false.

Exercise. C2: Challenge: Splitting an inconsistency, used later

Let Γ be a list and φ a statement.

(a) Prove that if Γ, φ and $\Gamma, \neg\varphi$ are both inconsistent, then so is Γ.

(b) Deduce: if Γ is consistent, then at least one of Γ, φ and $\Gamma, \neg\varphi$ is consistent.

(c) Does (b) say which one?

Solution.

(a) Since Γ, φ is inconsistent, Proposition 3.2.2 (a) gives $\Gamma \vdash \neg\varphi$. Since $\Gamma, \neg\varphi$ is inconsistent, part (b) gives $\Gamma \vdash \varphi$. So (iii) of Proposition 3.2.1 holds with χ the statement φ, and Γ is inconsistent.

(b) This is the contrapositive of (a): if both lists were inconsistent, Γ would be.

(c) No. It says that one of the two lists is consistent and gives no way to tell which. Deciding that would need a way to show a list consistent, which this chapter does not have.

Section 4.10 (optional) adds statements to a consistent list one at a time, in exactly this way.

Marking scheme. (10 points)

- 4: (a) obtains $\Gamma \vdash \neg\varphi$ by part (a) of the proposition or $\Gamma \vdash \varphi$ by part (b), and combines it with the other inconsistency, directly (by Cut, Lemma 2.3.3 (b)) or by the other part.
- 3: (a) concludes by Proposition 3.2.1, by Lemma 2.3.3 (c) with one ¬E, or by Cut (Lemma 2.3.3 (b)).
- 2: (b) as the contrapositive.
- 1: (c) "no".

Fatal error: (b) claimed to show that a particular list is consistent.

3.3 From Argument to Derivation

Most searches in Sections 3.1 and 3.2 started from a sequent: its letters, its premises and its conclusion were given. Mathematics starts from a paragraph. Its premises are scattered among words such as "since" and "so"; some of them are not written at all; and its letters must be chosen. This section turns paragraphs into sequents and derives them. Then it goes the other way, and writes a proof beside the derivation it mirrors. A last page collects the sequents that searches keep needing.

3.3.1 One key for the section

Section 1.1, Exercise B5, fixed one key for an integer n, and this section keeps it: E: "n is even", P: "n is prime", T: "n is a multiple of 3", G: "$n > 10$" (Exercise (B5: One key, sixteen sentences, used later)). A **prime** is an integer greater than 1 whose only positive divisors are 1 and itself. Three facts about integers will serve as premises:

- **F1** $(P \wedge E) \to \neg G$: an even prime is at most 10. The only even prime is 2, since an even $n > 2$ has the divisor 2.
- **F2** $(P \wedge T) \to \neg G$: a prime multiple of 3 is at most 10. The only one is 3, since a multiple of 3 greater than 3 has the divisor 3.
- **F3** $(E \wedge T) \to \neg P$: no even multiple of 3 is prime. Such an integer is a multiple of 6. A positive multiple of 6 has the divisor 2 and is greater than 2, and 0 and the negative integers are not prime.

Each fact is true for every integer n. An argument may use any of them as a premise, and a

derivation lists the ones it uses.

A translation is checked by reading it back, letter by letter, as in Section 1.1: F3 reads back as "if n is even and a multiple of 3, then n is not prime".

Quick check. On the key, translate "n is an odd prime greater than 10". Then read back $G \to \neg(P \wedge E)$.

Solution. $(\neg E \wedge P) \wedge G$. And $G \to \neg(P \wedge E)$ reads back as "if $n > 10$, then n is not an even prime".

3.3.2 Extracting an argument

To turn a paragraph into a sequent, work through five steps.

1. *Find the conclusion.* Words such as *so*, *hence*, *therefore*, *thus* and *it follows that* introduce it.
2. *Find the stated premises.* Words such as *since*, *because*, *for* and *given that* introduce them. A supposition ("suppose ...", "let n be ...") is a premise, or the assumption of a box.
3. *Name the unstated facts.* Wherever a step does not follow by a rule from the lines before it, write the fact that would make it follow, as an implication, and check that the fact is true. Supply the weakest fact that closes the gap: a stronger one may prove more than the paragraph claims, or be false. (Compare Section 1.3, Exercise B5 (d), where a missing implication was written and checked.)
4. *Translate on a key*, with Section 1.1's phrase book.
5. *Derive the sequent, or break it.*

Steps 1–4 are reading. Step 5 is the rest of this book.

Example. An argument about a prime, extracted

Extract the argument of this paragraph and derive it.

> Let n be an integer, and suppose that n is a prime greater than 10. Then n is odd, since the only even prime is 2. Nor is n a multiple of 3, because the only prime that is a multiple of 3 is 3 itself. So n is neither even nor a multiple of 3.

Solution. We take the five steps in order.

1. The conclusion follows "So": n is neither even nor a multiple of 3, that is, $\neg E \wedge \neg T$ ("neither ... nor" in the phrase book).
2. The stated premise is the supposition, "n is a prime greater than 10": $P \wedge G$, since "a prime greater than 10" is a conjunction. "n is odd" and "nor is n a multiple of 3" are not premises. They are steps, each with a reason after "since" or "because".
3. The reasons are facts about numbers, but not in a form a rule can use. "The only even prime is 2" is used as "an even prime is not greater than 10", which is F1. The second reason is used as F2.

4. The sequent is

$$P \wedge G,\ (P \wedge E) \to \neg G,\ (P \wedge T) \to \neg G \vdash \neg E \wedge \neg T.$$

5. We read the derivation off the goal. It is a conjunction of two negations: two ¬I boxes joined by ∧I, as in Section 3.1.

In the first picture, the premise is taken apart and the two conjuncts stand as subgoals. In the second, each subgoal has its box, with goal ⊥.

1	$P \wedge G$	
2	$(P \wedge E) \to \neg G$	
3	$(P \wedge T) \to \neg G$	
4	P	∧E 1
5	G	∧E 1
	⋮	
10	$\neg E$	
	⋮	
15	$\neg T$	
16	$\neg E \wedge \neg T$	∧I 10, 15

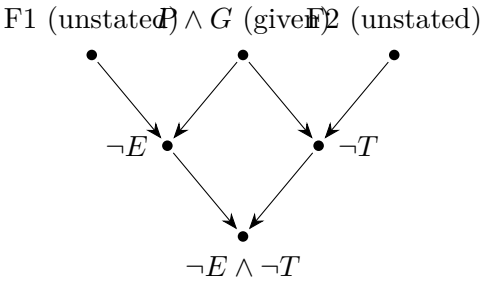
1	$P \wedge G$	
2	$(P \wedge E) \to \neg G$	
3	$(P \wedge T) \to \neg G$	
4	P	∧E 1
5	G	∧E 1
6	E	asm, for ¬I
	⋮	
9	⊥	
10	$\neg E$	¬I 6–9
11	T	asm, for ¬I
	⋮	
14	⊥	
15	$\neg T$	¬I 11–14
16	$\neg E \wedge \neg T$	∧I 10, 15

Inside each box, the fact fires once the box's assumption is joined to P, and its conclusion $\neg G$ clashes with line 5.

1	$P \wedge G$	
2	$(P \wedge E) \to \neg G$	
3	$(P \wedge T) \to \neg G$	
4	P	$\wedge$E 1
5	G	$\wedge$E 1
6	$\quad E$	asm, for $\neg$I
7	$\quad P \wedge E$	$\wedge$I 4, 6
8	$\quad \neg G$	$\to$E 2, 7
9	$\quad \bot$	$\neg$E 5, 8
10	$\neg E$	$\neg$I 6–9
11	$\quad T$	asm, for $\neg$I
12	$\quad P \wedge T$	$\wedge$I 4, 11
13	$\quad \neg G$	$\to$E 3, 12
14	$\quad \bot$	$\neg$E 5, 13
15	$\neg T$	$\neg$I 11–14
16	$\neg E \wedge \neg T$	$\wedge$I 10, 15

Lines 4–5 need nothing new, so they are written before the boxes (Section 3.1). Had the supposition been kept as a supposition, the claim would be $(P \wedge E) \to \neg G,\ (P \wedge T) \to \neg G \vdash (P \wedge G) \to (\neg E \wedge \neg T)$: the same lines inside one more box, closed by $\to$I.

The argument can also be drawn as a map: each statement a dot, each step an arrow into what it gives. The unstated facts are the dots nobody wrote. The map shows what supports what; it is not a derivation, and it leaves out the boxes.



Quick check. Which line of the derivation in Example (An argument about a prime, extracted) is the paragraph's sentence "Then n is odd", and which lines does the paragraph leave silent?

Solution. Line 10. The paragraph is silent on lines 4–5 and on the box, lines 6–9, that shows it, apart from the reason F1 behind line 8.

Warning (Not allowed: extracting an argument). Each entry names the step of the method it breaks.

- "Writing 'n is odd' among the premises": a step taken as a premise. The derivation then proves nothing about whether n is odd (step 2).
- "Reading 'n is odd, since n is a prime greater than 10' with 'n is odd' as the premise": *since* introduces the premise, not the conclusion (steps 1–2).
- "Supplying 'every prime is odd' for the first step": the supplied fact must be true, and this one fails at $n = 2$ (step 3).
- "Supplying $\neg E \wedge \neg T$ itself as the unstated fact": the derivation is legal and proves nothing. Supply the weakest fact that closes the gap (step 3).
- "Translating 'n is a prime greater than 10' as $P \to G$": a noun with an adjective is a conjunction, $P \wedge G$, as "an even prime" was $E \wedge P$ in Section 1.1 (step 4).

3.3.3 Two columns

Now the other direction: a proof is written, and its derivation is wanted. Write the proof in the left column and the derivation in the right, each sentence on the row of the lines it says. A line that no sentence says is a **silent step**. A sentence that matches no line is a gap in the proof, or a step this book's rules cannot take.

Example. A proof and its derivation, side by side

Let n be an integer. Claim: if n is even and prime, then n is not a multiple of 3.

> *Proof.* Suppose n is even and prime. If n were a multiple of 3, it would be an even multiple of 3, and no even multiple of 3 is prime; but n is prime. So n is not a multiple of 3.

Write the derivation the proof mirrors, in two columns.

Solution. The proof uses one fact, F3, and the claim is an implication. The sequent is $(E \wedge T) \to \neg P \vdash (E \wedge P) \to \neg T$. The goal gives a box on $E \wedge P$ (left). Inside it the goal $\neg T$ is a negation, so a $\neg$I box on T opens, after the assumption has been taken apart (right).

1	$(E \wedge T) \to \neg P$	
2	$\quad E \wedge P$	asm, for $\to$I
	$\quad \vdots$	
9	$\quad \neg T$	
10	$(E \wedge P) \to \neg T$	$\to$I 2–9

1	$(E \wedge T) \to \neg P$	
2	$\quad E \wedge P$	asm, for $\to$I
3	$\quad E$	$\wedge$E 2
4	$\quad P$	$\wedge$E 2
5	$\quad\quad T$	asm, for $\neg$I
	$\quad\quad \vdots$	
8	$\quad\quad \bot$	
9	$\quad \neg T$	$\neg$I 5–8
10	$(E \wedge P) \to \neg T$	$\to$I 2–9

The two columns finish the derivation.

The proof	The derivation
"Suppose n is even and prime."	2 $E \wedge P$, asm, for $\to$I
(silent)	3 E, 4 P, $\wedge$E 2
"If n were a multiple of 3,"	5 T, asm, for $\neg$I
"it would be an even multiple of 3,"	6 $E \wedge T$, $\wedge$I 3, 5
"and no even multiple of 3 is prime;"	1 $(E \wedge T) \to \neg P$, F3; 7 $\neg P$, $\to$E 1, 6
"but n is prime."	8 $\bot$, $\neg$E 4, 7
"So n is not a multiple of 3."	9 $\neg T$, $\neg$I 5–8
(the claim)	10 $(E \wedge P) \to \neg T$, $\to$I 2–9

The finished derivation, read down the right column:

1	$(E \wedge T) \to \neg P$	
2	$\quad E \wedge P$	asm, for $\to$I
3	$\quad E$	$\wedge$E 2
4	$\quad P$	$\wedge$E 2
5	$\quad\quad T$	asm, for $\neg$I
6	$\quad\quad E \wedge T$	$\wedge$I 3, 5
7	$\quad\quad \neg P$	$\to$E 1, 6
8	$\quad\quad \bot$	$\neg$E 4, 7
9	$\quad \neg T$	$\neg$I 5–8
10	$(E \wedge P) \to \neg T$	$\to$I 2–9

The proof says lines 2 and 5–9, and line 1 only inside a clause. It is silent on lines 3–4, taking the conjunction apart, and on line 10, since closing the outer box is the claim itself. Its "Suppose" is line 2's "asm, for $\to$I", and its "If n were" is line 5's "asm, for $\neg$I". The proof hides the fact F3 in a clause; the derivation lists it.

Quick check. If the claim in Example (A proof and its derivation, side by side) were "if n is even and prime, then n is at most 10", which fact would replace F3, and what would happen to the box of line 5?

Solution. F1 replaces F3. The box of line 5 disappears: the goal $\neg G$ is F1's conclusion, so after $\wedge$E and $\wedge$I give $P \wedge E$, one $\to$E line reaches it.

3.3.4 A library of standard sequents

Chapters 1 and 2 and Section 3.2 proved sequents that searches keep needing. Here they are grouped by what they are for: a searcher with a resource or a goal of a given shape looks up its row. Each row is a derived rule (Definition 2.3.1). It is applied to **whole** available lines, never to a part of a line, and it is cited by the name in its middle column. "The rules in one place" in Section 3.1 lists the primitive rules and most of these derived rules on one card; this page regroups the derived rules, with those of Section 3.2, by use. The last group is different: it holds facts about derivations, which change the task rather than add a line.

Sequent	In a justification	Proved in
To use an implication		
$P \to Q,\ \neg Q \vdash \neg P$	MT	Proposition 2.3.1
$P \to Q,\ Q \to R \vdash P \to R$	(implication is transitive)	Proposition 1.4.1
$(P \wedge Q) \to R \dashv\vdash P \to (Q \to R)$	(curry)	Proposition 1.4.2
$P \to Q \vdash \neg P \vee Q$	(implication as disjunction)	Proposition 2.5.6 (b)
To use a disjunction		
$P \vee Q,\ \neg P \vdash Q$; $P \vee Q,\ \neg Q \vdash P$	DS	Proposition 2.3.4; Exercise (B1: The other disjunctive syllogism, used later) (a)
$(P \to R) \wedge (Q \to R) \dashv\vdash (P \vee Q) \to R$	(cases on a hypothesis)	Example (Cases on a hypothesis are two implications); Example (A disjunctive hypothesis splits into two implications)
$(P \to R) \wedge (Q \to S),\ P \vee Q \vdash R \vee S$	(constructive dilemma)	Example (Constructive dilemma)
$(P \to R) \wedge (Q \to S),\ \neg R \vee \neg S \vdash \neg P \vee \neg Q$	(destructive dilemma)	Exercise (B4: Dilemmas as derived rules, used later) (b)
$\neg(P \wedge Q) \vdash P \to \neg Q$	(*modus ponendo tollens*)	Exercise (B4: Dilemmas as derived rules, used later) (c)
To use a negation or a clash		
$P,\ \neg P \vdash Q$	(ex falso)	Proposition 2.1.1
$\neg P \vdash P \to Q$	VI	Proposition 2.1.2
$\neg\neg P \vdash P$	$\neg\neg$E	Proposition 2.5.1 (b)
$\neg(P \vee Q) \vdash \neg P \wedge \neg Q$; $\neg(P \wedge Q) \vdash \neg P \vee \neg Q$	DM	Proposition 2.5.5 (a), (d)
To reach a negation		
$P \to Q \vdash \neg Q \to \neg P$	CP	Proposition 2.3.2
$\vdash (P \to (Q \wedge \neg Q)) \to \neg P$	(implies a contradiction)	Example (A statement that implies a contradiction)
$\neg P \wedge \neg Q \vdash \neg(P \vee Q)$; $\neg P \vee \neg Q \vdash \neg(P \wedge Q)$	DM	Proposition 2.5.5 (b), (c)
$P \vdash \neg\neg P$	$\neg\neg$I	Proposition 2.5.1 (a)
To reach an implication		
$\neg P \vee Q \vdash P \to Q$	(implication as disjunction)	Proposition 2.5.6 (a)

Sequent	In a justification	Proved in
$\neg Q \to \neg P \vdash P \to Q$	(CP, converse)	Proposition 2.5.2
To reach a disjunction		
$\vdash P \vee \neg P$	LEM	Proposition 2.5.3
$(P \wedge \neg Q) \to R \dashv\vdash P \to (Q \vee R)$	(refute one side)	Example (Proving a disjunction by refuting one side)
To rearrange		
$P \wedge Q \vdash Q \wedge P$; $P \vee Q \vdash Q \vee P$	(and commutes); (or commutes)	Proposition 1.2.1; Example (First derivations with "or")
$(P \wedge Q) \wedge R \vdash P \wedge (Q \wedge R)$; $P \vee (Q \vee R) \vdash (P \vee Q) \vee R$	(and regroups); (or regroups)	Proposition 1.2.2; Example (Regrouping a disjunction)
$P \wedge (Q \vee R) \dashv\vdash (P \wedge Q) \vee (P \wedge R)$	(distributivity)	Proposition 1.6.1
$P \leftrightarrow Q \vdash Q \leftrightarrow P$	(iff symmetric)	Example (An iff read both ways)
$P \leftrightarrow Q \dashv\vdash (P \to Q) \wedge (Q \to P)$	(iff as two implications)	Proposition 2.4.1
$P \to Q,\ Q \to R,\ R \to P \vdash P \leftrightarrow Q$	(equivalence cycle)	Proposition 2.4.6
To change the task (facts about derivations, not lines)		
more premises never hurt; a derived statement may be used as a premise	(weakening, cut)	Lemma 2.3.3
$\Gamma \vdash \varphi$ exactly when $\Gamma, \neg\varphi$ is inconsistent	—	Proposition 3.2.2
an inconsistent list derives every statement	—	Proposition 3.2.1

Rows proved in exercises (the second form of DS; the destructive dilemma and *modus ponendo tollens* rows, Exercise (B4: Dilemmas as derived rules, used later) (b), (c)) may be used by exercises, as in Section 3.1; the text's own derivations use only rows proved in the text.

Quick check. Which row would you look up first? (i) The goal is $\neg(P \vee Q)$. (ii) A resource is $\neg(P \wedge Q)$, and the goal is $\neg P \vee \neg Q$. (iii) The resources include $P \to Q$ and $\neg Q$.

Solution.

(i) DM (b), under "To reach a negation". (ii) DM (d), which gives the goal in one line. (iii) MT.

Remark (Looking ahead). We show that something is derivable by exhibiting a derivation; how could we ever show that something is not? Chapter 4 gives each statement a truth value and proves that the rules never lead from true premises to a false conclusion. Then one choice of truth values, such as P false and Q true for the failed search of Section 3.1,

proves that no derivation exists.

3.3.5 Exercises

3.3.5.1 A. Check your understanding

Exercise. A1

(a) In "Since n^2 is even, n is even", which is the premise and which the conclusion?

(b) Find the conclusion: "n is odd, for n is prime and greater than 10."

(c) Name the unstated fact: "n is a multiple of 6, so n is even."

(d) On the key of this section, translate "n is an odd prime at most 10."

(e) In Example (A proof and its derivation, side by side), which sentence of the proof is line 5, and what does line 5's annotation read?

Solution.

(a) The premise is "n^2 is even"; the conclusion is "n is even".

(b) "n is odd". *For* introduces the premise.

(c) Every multiple of 6 is even: if $6 \mid n$, then $2 \mid n$.

(d) $(\neg E \wedge P) \wedge \neg G$.

(e) "If n were a multiple of 3"; the annotation reads "asm, for ¬I".

3.3.5.2 B. Practice

Exercise. B1: Extract, then derive

For each paragraph, find the conclusion, the stated premises and any unstated fact; choose letters; write the sequent; derive it.

(a) Let x be a real number with $x^2 \leq x$. If x were greater than 1, then $x^2 = x \cdot x$ would be greater than x. If x were negative, x^2 would be positive, hence greater than x. So x is at least 0 and at most 1.

(b) Let n be an integer. Since n^2 is odd, n is odd.

(c) Let a and b be real numbers with $ab = 0$. If $a \neq 0$, we may divide by a and find $b = 0$. So $a = 0$ or $b = 0$.

Solution.

(a) Take A: "$x > 1$", B: "$x < 0$", C: "$x^2 > x$". The paragraph states two facts, each with its reason: $A \to C$ and $B \to C$. The premise "$x^2 \leq x$" is $\neg C$, by the convention of Section 2.1. The conclusion "at least 0 and at most 1" is $\neg B \wedge \neg A$, by the same convention. The sequent is $A \to C,\ B \to C,\ \neg C \vdash \neg B \wedge \neg A$.

1	$A \to C$	
2	$B \to C$	
3	$\neg C$	
4	$\neg A$	MT 1, 3
5	$\neg B$	MT 2, 3
6	$\neg B \wedge \neg A$	$\wedge$I 5, 4

Nothing is unstated here except the order convention.

(b) Take E: "n is even", S: "n^2 is even". The premise is $\neg S$ and the conclusion $\neg E$. The unstated fact is $E \to S$, "the square of an even integer is even", which is true since $(2k)^2 = 4k^2 = 2(2k^2)$. The sequent is $E \to S,\ \neg S \vdash \neg E$.

1	$E \to S$	
2	$\neg S$	
3	$\neg E$	MT 1, 2

The tempting wrong supply is "the square of an odd integer is odd", $\neg E \to \neg S$. It is true, since $(2k+1)^2 = 2(2k^2+2k)+1$, but with the premise $\neg S$ it gives nothing: concluding $\neg E$ from it would be affirming the consequent (Section 1.3).

(c) Take A: "$a = 0$", B: "$b = 0$", Z: "$ab = 0$". The stated fact is $(Z \wedge \neg A) \to B$, dividing by a: $b = a^{-1}(ab) = 0$. The premise is Z and the conclusion $A \vee B$. Neither side of the goal is in hand, so we split on A by LEM, Section 3.1's step 5, as in Example (Proving a disjunction by refuting one side).

1	$(Z \wedge \neg A) \to B$	
2	Z	
3	$A \vee \neg A$	LEM
4	$\quad A$	asm, for $\vee$E
5	$\quad A \vee B$	$\vee$I 4
6	$\quad \neg A$	asm, for $\vee$E
7	$\quad Z \wedge \neg A$	$\wedge$I 2, 6
8	$\quad B$	$\to$E 1, 7
9	$\quad A \vee B$	$\vee$I 8
10	$A \vee B$	$\vee$E 3, 4–5, 6–9

Exercise. B2: Two columns

Let n be an integer. Claim: if n is prime and $n > 10$, then n is not a multiple of 6.

Proof. Suppose n is prime and $n > 10$. If n were a multiple of 6, it would be an even multiple of 3, and no even multiple of 3 is prime; but n is prime. So n is not a multiple of 6.

(a) On the key, "n is a multiple of 6" is $E \wedge T$. The derivation the proof mirrors has reached the stage below. *Name the move, then complete it.*

1	$(E \wedge T) \to \neg P$	
2	$\quad P \wedge G$	asm, for $\to$I
	$\quad \vdots$	
7	$\quad \neg(E \wedge T)$	
8	$(P \wedge G) \to \neg(E \wedge T)$	$\to$I 2–7

(b) Write the two-column table.

(c) Which hypothesis does the derivation never cite? What claim does that show, and is it true?

Solution.

(a) The goal $\neg(E \wedge T)$ is a negation, so step 2 of the method opens a $\neg$I box on $E \wedge T$. F3 fires on it, and the clash partner P comes from line 2.

1	$(E \wedge T) \to \neg P$	
2	$\quad P \wedge G$	asm, for $\to$I
3	$\quad P$	$\wedge$E 2
4	$\quad\quad E \wedge T$	asm, for $\neg$I
5	$\quad\quad \neg P$	$\to$E 1, 4
6	$\quad\quad \bot$	$\neg$E 3, 5
7	$\quad \neg(E \wedge T)$	$\neg$I 4–6
8	$(P \wedge G) \to \neg(E \wedge T)$	$\to$I 2–7

(b) The two columns:

The proof	The derivation
"Suppose n is prime and $n > 10$."	2
(silent)	3
"If n were a multiple of 6,"	4
"it would be an even multiple of 3, and no even multiple of 3 is prime;"	1 (F3), 5
"but n is prime."	6
"So n is not a multiple of 6."	7
(the claim)	8

Line 4 already is $E \wedge T$, so "an even multiple of 3" is a rereading, not a step.

(c) G, that is, $n > 10$, is never cited: no line uses it. The same lines with P in place of $P \wedge G$ as the assumption of the outer box (line 3 dropped), closed by $\to$I, would derive $(E \wedge T) \to \neg P \vdash P \to \neg(E \wedge T)$: no prime is a multiple of 6. That is true for every integer, by F3. The hypothesis $n > 10$ was idle.

Exercise. B3: Fifteen items

Items 1–5: derive each sequent, using a row of the library where it shortens the derivation, and name the row; their letters are free, not the key. Items 6–10: translate on the key, using the facts F1–F3 as premises where needed, and derive. Items 11–15: derive or break, as in Example (Two claims, one connective apart). The facts may be used as premises. A break is an integer n at which the premises are true and the conclusion false; since F1–F3 hold at every n, such an n makes the whole counterexample set true.

Derive, with the library.

1. $\neg(P \vee Q) \vdash P \to R$.
2. $P \vee Q,\ P \to R,\ \neg R \vdash Q \wedge \neg P$.
3. $\neg(P \wedge \neg Q) \vdash P \to Q$.
4. $(P \vee Q) \to R \vdash \neg R \to (\neg P \wedge \neg Q)$.
5. $\neg P \to Q,\ \neg Q \vdash P \wedge (\neg Q \vee R)$.

Translate on the key, then derive.

6. n is an even multiple of 3; so n is not prime, and n is not a prime greater than 10.
7. n is a prime at most 10, or n is odd; so if n is even, n is at most 10.
8. n is a prime multiple of 3; so n is at most 10, and n is odd.
9. n is prime; n is not a multiple of 3 or n is at most 10; and n is a multiple of 3. So n is at most 10.
10. If n is greater than 10, then n is not an even prime.

Derive or break.

11. n is even but not a multiple of 3; so n is not prime.
12. n is an even prime; so n is at most 10.
13. n is odd and greater than 10; so n is prime.
14. n is odd; so n is not an even prime.
15. n is not prime; so n is even or a multiple of 3.

Solution. **1.** DM (a), then VI.

1	$\neg(P \vee Q)$	
2	$\neg P \wedge \neg Q$	DM 1
3	$\neg P$	$\wedge$E 2
4	$P \to R$	VI 3

2. MT, then DS.

1	$P \vee Q$	
2	$P \to R$	
3	$\neg R$	
4	$\neg P$	MT 2, 3
5	Q	DS 1, 4
6	$Q \wedge \neg P$	$\wedge$I 5, 4

3. *Modus ponendo tollens*, with $\neg Q$ for Q, turns the premise into $P \to \neg\neg Q$. Inside a box on P, $\to$E and $\neg\neg$E reach Q.

1	$\neg(P \wedge \neg Q)$	
2	$P \to \neg\neg Q$	modus ponendo tollens 1
3	$\quad P$	asm, for $\to$I
4	$\quad \neg\neg Q$	$\to$E 2, 3
5	$\quad Q$	$\neg\neg$E 4
6	$P \to Q$	$\to$I 3–5

Without the library, the goal Q inside the box is a letter, so RAA on $\neg Q$, which clashes with the premise once $P \wedge \neg Q$ is built: seven lines.

1	$\neg(P \wedge \neg Q)$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad \neg Q$	asm, for RAA
4	$\quad\quad P \wedge \neg Q$	$\wedge$I 2, 3
5	$\quad\quad \bot$	$\neg$E 4, 1
6	$\quad Q$	RAA 3–5
7	$P \to Q$	$\to$I 2–6

The other library route (DM (d) on the premise, then $\neg\neg$I, DS and $\neg\neg$E inside the box) is no shorter than the seven lines.

4. MT, then DM (a), inside one box.

1	$(P \vee Q) \to R$	
2	$\quad \neg R$	asm, for $\to$I
3	$\quad \neg(P \vee Q)$	MT 1, 2
4	$\quad \neg P \wedge \neg Q$	DM 3
5	$\neg R \to (\neg P \wedge \neg Q)$	$\to$I 2–4

5. MT, then $\neg\neg$E.

1	$\neg P \to Q$	
2	$\neg Q$	
3	$\neg\neg P$	MT 1, 2
4	P	$\neg\neg$E 3
5	$\neg Q \vee R$	$\vee$I 2
6	$P \wedge (\neg Q \vee R)$	$\wedge$I 4, 5

6. $E \wedge T,\ (E \wedge T) \to \neg P \vdash \neg P \wedge \neg(P \wedge G)$, with F3.

1	$E \wedge T$	
2	$(E \wedge T) \to \neg P$	
3	$\neg P$	$\to$E 2, 1
4	$\quad P \wedge G$	asm, for $\neg$I
5	$\quad P$	$\wedge$E 4
6	$\quad \bot$	$\neg$E 5, 3
7	$\neg(P \wedge G)$	$\neg$I 4–6
8	$\neg P \wedge \neg(P \wedge G)$	$\wedge$I 3, 7

7. $(P \wedge \neg G) \vee \neg E \vdash E \to \neg G$, with no fact needed.

1	$(P \wedge \neg G) \vee \neg E$	
2	$\quad E$	asm, for $\to$I
3	$\quad\quad P \wedge \neg G$	asm, for $\vee$E
4	$\quad\quad \neg G$	$\wedge$E 3
5	$\quad\quad \neg E$	asm, for $\vee$E
6	$\quad\quad \bot$	$\neg$E 2, 5
7	$\quad\quad \neg G$	$\bot$E 6
8	$\quad \neg G$	$\vee$E 1, 3–4, 5–7
9	$E \to \neg G$	$\to$I 2–8

8. $P \wedge T,\ (P \wedge T) \to \neg G,\ (E \wedge T) \to \neg P \vdash \neg G \wedge \neg E$, with F2 and F3.

1	$P \wedge T$	
2	$(P \wedge T) \to \neg G$	
3	$(E \wedge T) \to \neg P$	
4	$\neg G$	$\to$E 2, 1
5	P	$\wedge$E 1
6	T	$\wedge$E 1
7	$\quad E$	asm, for $\neg$I
8	$\quad E \wedge T$	$\wedge$I 7, 6
9	$\quad \neg P$	$\to$E 3, 8
10	$\quad \bot$	$\neg$E 5, 9
11	$\neg E$	$\neg$I 7–10
12	$\neg G \wedge \neg E$	$\wedge$I 4, 11

9. $P,\ \neg T \vee \neg G,\ T \vdash \neg G$.

1	P	
2	$\neg T \vee \neg G$	
3	T	
4	$\neg\neg T$	$\neg\neg$I 3
5	$\neg G$	DS 2, 4

The premise P is not needed.

10. $(P \wedge E) \to \neg G \vdash G \to \neg(P \wedge E)$, with F1.

1	$(P \wedge E) \to \neg G$	
2	$\quad G$	asm, for $\to$I
3	$\quad\quad P \wedge E$	asm, for $\neg$I
4	$\quad\quad \neg G$	$\to$E 1, 3
5	$\quad\quad \bot$	$\neg$E 2, 4
6	$\quad \neg(P \wedge E)$	$\neg$I 3–5
7	$G \to \neg(P \wedge E)$	$\to$I 2–6

With the library, lines 3–6 shrink to two: 3 $\neg\neg G$, $\neg\neg$I 2; 4 $\neg(P \wedge E)$, MT 1, 3.

11. Breaks. The claim is $E \wedge \neg T \vdash \neg P$, and at $n = 2$ the integer is even, not a multiple of 3, and prime.

12. Derives: $P \wedge E,\ (P \wedge E) \to \neg G \vdash \neg G$, with F1.

1	$P \wedge E$	
2	$(P \wedge E) \to \neg G$	
3	$\neg G$	$\to$E 2, 1

13. Breaks. The claim is $\neg E \wedge G \vdash P$, and at $n = 15$ the integer is odd, greater than 10, and not prime.

14. Derives: $\neg E \vdash \neg(P \wedge E)$. No fact is needed.

1	$\neg E$	
2	$\quad P \wedge E$	asm, for $\neg$I
3	$\quad E$	$\wedge$E 2
4	$\quad \bot$	$\neg$E 3, 1
5	$\neg(P \wedge E)$	$\neg$I 2–4

15. Breaks. The claim is $\neg P \vdash E \vee T$, and at $n = 25$ the integer is not prime, odd, and not a multiple of 3.

Items 11 and 12, and items 13 and 14, pair a break with a derivation on nearby words; item 15 tries to go back from "not prime" to "even or a multiple of 3", which F3 does not give, and fails.

Exercise. B4: Put the proof in order

Let a and b be real numbers. Claim: if $a + b > 2$, then $a > 1$ or $b > 1$.

(i) This contradicts $a + b > 2$.
(ii) Suppose $a + b > 2$.
(iii) Then $a > 1$, since $a + b > 2$.
(iv) Adding, $a + b \leq 2$.
(v) Hence $a > 1$ or $b > 1$.
(vi) Suppose, for a contradiction, that neither $a > 1$ nor $b > 1$.
(vii) Suppose $a > 1$ or $b > 1$.
(viii) Then $a \leq 1$ and $b \leq 1$.

(a) Put six of these sentences in an order that proves the claim; two do not belong.
(b) Match each sentence of your order to a line of the derivation below, with A: "$a > 1$", B: "$b > 1$", S: "$a + b > 2$".

1	$(\neg A \wedge \neg B) \to \neg S$	
2	$\quad S$	asm, for $\to$I
3	$\quad\quad \neg(A \vee B)$	asm, for RAA
4	$\quad\quad \neg A \wedge \neg B$	DM 3
5	$\quad\quad \neg S$	$\to$E 1, 4
6	$\quad\quad \bot$	$\neg$E 2, 5
7	$\quad A \vee B$	RAA 3–6
8	$S \to (A \vee B)$	$\to$I 2–7

Solution.

(a) (ii), (vi), (viii), (iv), (i), (v). No other order is admissible: the supposition (ii) must open first, since the claim is an implication, the RAA supposition (vi) inside it, and each later sentence uses the one before it. The two distractors:

- (i) does not follow: $a = 0$, $b = 3$ has $a + b > 2$ and $a \leq 1$.
- (i) assumes the conclusion, a box opened on the goal (Section 3.1's warning "Do not assume what you want to prove").

(g) (i) is line 2, "asm, for $\to$I". (vi) is line 3, "asm, for RAA". (viii) is line 4, by DM and the order convention of Section 2.1 ($\neg(a > 1)$ is $a \leq 1$). (iv) is line 5, with the fact on line 1: adding the two inequalities. (i) is line 6, and (v) is line 7. The claim is line 8. The fact on line 1 is the one step of school algebra: $a \leq 1$ and $b \leq 1$ give $a + b \leq 2$.

3.3.5.3 C. Going deeper

Exercise. C1: Exit check: from a paragraph to a derivation

Let n be an integer greater than 10, and suppose n is even or a multiple of 3. Then n is not prime. Indeed, suppose n were prime. If n is even, n is an even prime, and every even prime is at most 10. If n is a multiple of 3, n is a prime multiple of 3, and every such number is at most 10. Either way n is at most 10, which is false. So n is not prime.

(a) Name the conclusion, the stated premises, and the facts the paragraph gives as reasons, each on the key.

(b) Write the sequent.

(c) Derive it.

(d) Write the two-column table.

(e) Is the premise "$n > 10$" needed? Break the claim without it.

Solution.

(a) The conclusion is $\neg P$. The premises are G and $E \vee T$. The facts are F1 ("every even prime is at most 10") and F2.

(b) $G,\ E \vee T,\ (P \wedge E) \to \neg G,\ (P \wedge T) \to \neg G \vdash \neg P$.

(c) The goal is a negation: a ¬I box on P. Inside it, the resource $E \vee T$ is split, each case reaching $\neg G$ by its fact, and $\neg G$ clashes with G outside the cases.

1	G	
2	$E \vee T$	
3	$(P \wedge E) \to \neg G$	
4	$(P \wedge T) \to \neg G$	
5	$\quad P$	asm, for ¬I
6	$\quad\quad E$	asm, for ∨E
7	$\quad\quad P \wedge E$	∧I 5, 6
8	$\quad\quad \neg G$	→E 3, 7
9	$\quad\quad T$	asm, for ∨E
10	$\quad\quad P \wedge T$	∧I 5, 9
11	$\quad\quad \neg G$	→E 4, 10
12	$\quad \neg G$	∨E 2, 6–8, 9–11
13	$\quad \bot$	¬E 1, 12
14	$\neg P$	¬I 5–13

(d) The two columns:

The proof	The derivation
"Let n ... greater than 10, and suppose n is even or a multiple of 3."	1, 2
"Then n is not prime."	the goal, line 14
"suppose n were prime."	5 (asm, for ¬I)
"If n is even, ... at most 10."	6–8, with 3
"If n is a multiple of 3, ... at most 10."	9–11, with 4
"Either way n is at most 10,"	12
"which is false."	13
"So n is not prime."	14

(e) It is needed. Without G, the claim $E \vee T,\ (P \wedge E) \to \neg G,\ (P \wedge T) \to \neg G \vdash \neg P$ breaks at $n = 2$: 2 is even, so $E \vee T$ holds; F1 and F2 hold, as at every n; and 2 is prime, so the negated conclusion $\neg\neg P$ is true too.

Marking scheme. (10 points)

- 2: (a)–(b), the sequent, with F1 and F2 named as the facts given as reasons, in implication form.
- 3: (c), a correct derivation; the model has the ¬I box on P, the two cases both ending in $\neg G$, and the clash with G outside the cases. In (d), the table is marked against the reader's own derivation.
- 2: (d), every sentence placed, the two "suppose"s matched to their annotations.
- 3: (e), "needed", with a break that makes every member of the new counterexample set true.

Fatal errors (0 for the step): $\neg P$ taken as a premise; the cases of ∨E ending in different statements; a "break" in (e) at an n where a premise fails. A derivation that splits $E \vee T$ first and opens a ¬I box in each case earns full marks for (c).

Exercise. C2: Challenge: A statement that promises anything

(a) Derive $C \leftrightarrow (C \to Q) \vdash Q$.

(b) Let Q be "$0 = 1$". What does (a) say about a statement C that is equivalent to "if C, then $0 = 1$"?

(c) Which rules does your derivation use? Compare Lemma 2.4.2.

Solution.

(a) The premise gives nothing until one side of the iff is on a line. The side $C \to Q$ is an implication, and its box can use the premise on its own assumption.

1	$C \leftrightarrow (C \to Q)$	
2	$\quad C$	asm, for →I
3	$\quad C \to Q$	↔E 1, 2
4	$\quad Q$	→E 3, 2
5	$C \to Q$	→I 2–4
6	C	↔E 1, 5
7	$C \to Q$	↔E 1, 6
8	Q	→E 7, 6

The shorter ending 7 Q, →E 5, 6 is a detour: line 5 was built by →I and is taken apart at once. Cutting it costs a line here, because line 5 is cited again at line 6, so it stays. This is the case Section 3.1 names: cutting a detour need not shorten a derivation, and the rule is advice, not a theorem.

(b) Read as an argument, the derivation proves $0 = 1$ from the condition on C; since $0 = 1$ is false, no statement can meet the condition. In this chapter's words, the list $C \leftrightarrow (C \to Q),\ \neg Q$ is inconsistent: add $\neg Q$ as a second premise (line 2, the later lines

shifting down by one) and a last line $\bot$, $\neg$E, citing Q and that premise.

(c) Only $\leftrightarrow$E, $\rightarrow$I and $\rightarrow$E: no rule for $\neg$ or $\bot$, and no RAA. Lemma 2.4.2 shows that no statement is equivalent to its own negation. This shows more: any statement equivalent to "if C, then Q" forces Q, whatever Q is. So when Q is refutable, as "$0 = 1$" is, no statement is equivalent to "if C, then Q", and the argument uses no "not". (Taking Q to be $\bot$ gives back Lemma 2.4.2, since $\neg C$ and $C \rightarrow \bot$ are interderivable (Section 2.2).) The trouble is self-reference, not negation. This is known as Curry's paradox, after the logician whose name is also on the law Proposition 1.4.2. Section 5.6 meets the same shape with "for every", in the barber and Russell's theorem.

Marking scheme. (10 points)

- 4: (a), any correct derivation; the model has the box on C (lines 2–4, with the assumption C cited twice), C recovered outside it by $\leftrightarrow$E (line 6), and Q reached outside every box. The 7-line ending $\rightarrow$E 5, 6 earns full marks; the detour remark is not required.
- 3: (b), the conclusion that no statement meets the condition, with the inconsistency.
- 3: (c), the rules of the reader's own derivation named correctly, and the point that a derivation using only $\leftrightarrow$E, $\rightarrow$I and $\rightarrow$E exists (the model), so no negation rule is needed.

Fatal error (0 for (a)): Q obtained inside the box and carried out of it.

3.3.5.4 R. Review

Exercise. R1

For real numbers a and b, let $\max(a, b) \coloneqq a$ and $\min(a, b) \coloneqq b$ if $a \geq b$, and $\max(a, b) \coloneqq b$ and $\min(a, b) \coloneqq a$ if $a < b$. Prove that $\max(a, b) + \min(a, b) = a + b$.

Solution. Since $a \geq b$ or $a < b$, and these two cases cover every pair of real numbers, we argue by cases ($\vee$E, Definition 1.6.1), as in Example (The absolute value of a product).

Case $a \geq b$. Then $\max(a, b) = a$ and $\min(a, b) = b$, so $\max(a, b) + \min(a, b) = a + b$.

Case $a < b$. Then $\max(a, b) = b$ and $\min(a, b) = a$, so $\max(a, b) + \min(a, b) = b + a = a + b$.

In either case the equation holds, which proves it. Both cases end in the same statement, the goal, as $\vee$E requires.

Exercise. R2

Let x be a real number. Write N for "$x < 0$", M for "$x > 1$", and S for "$x = 7$". Translate the argument below into a sequent in these letters, and derive it.

> x is neither negative nor greater than 1. Therefore x is not negative, and if x is greater than 1, then $x = 7$.

Solution. The phrase book (Section 1.1) gives "neither ... nor" as $\neg N \wedge \neg M$; equivalently (Proposition 2.5.5 (a), (b)) $\neg(N \vee M)$, "$N \vee M$ fails", which we use here as the premise. It is not $\neg N \vee M$. The conclusion is a conjunction whose second part is an implication. The

sequent (Definition 1.2.4) is

$$\neg(N \vee M) \vdash \neg N \wedge (M \to S).$$

The goal is a conjunction; the first conjunct is a negation (a ¬I box on N), and the second an implication (a box on M). In both boxes ∨I builds $N \vee M$, which clashes with the premise; in the second box ⊥E then gives S.

1	$\neg(N \vee M)$	
2	$\quad N$	asm, for ¬I
3	$\quad N \vee M$	∨I 2
4	$\quad \bot$	¬E 3, 1
5	$\neg N$	¬I 2–4
6	$\quad M$	asm, for →I
7	$\quad N \vee M$	∨I 6
8	$\quad \bot$	¬E 7, 1
9	$\quad S$	⊥E 8
10	$M \to S$	→I 6–9
11	$\neg N \wedge (M \to S)$	∧I 5, 10

Lines 2–5 are half of Proposition 2.5.5 (a), which turns $\neg(N \vee M)$ into the phrase book's $\neg N \wedge \neg M$; with $\neg N \wedge \neg M$ as the premise, ∧E shortens both boxes. The implication $M \to S$ is vacuously true, as in Proposition 2.1.2: the premise rules out M, and a promise never tested is never broken, whatever S says.

Exercise. R3

Is $(P \to Q) \to R \vdash P \to (Q \to R)$ derivable? Derive or break its converse, $P \to (Q \to R) \vdash (P \to Q) \to R$.

Solution. The first is derivable, by the search method of Section 3.1. The goal gives a box on P, then a box on Q, with the goal R. The premise gives R by →E once we produce $P \to Q$, a new goal: a box on P whose last line repeats Q.

1	$(P \to Q) \to R$	
2	$\quad P$	asm, for $\to$I
3	$\quad\quad Q$	asm, for $\to$I
4	$\quad\quad\quad P$	asm, for $\to$I
5	$\quad\quad\quad Q$	R 3
6	$\quad\quad P \to Q$	$\to$I 4–5
7	$\quad\quad R$	$\to$E 1, 6
8	$\quad Q \to R$	$\to$I 3–7
9	$P \to (Q \to R)$	$\to$I 2–8

The repetition at line 5 puts Q inside the box of line 4, as $\to$I requires (clause (D3) of Definition 1.4.2). Two look-alike tools do not apply. Proposition 1.4.2 moves a **conjunction** in and out of a hypothesis, $(P \wedge Q) \to R \dashv\vdash P \to (Q \to R)$, and the premise here has $P \to Q$, not $P \wedge Q$, in that place. And Proposition 1.4.1 chains two implications, which we do not have.

The converse breaks: with P and R false (and Q either way), the premise $P \to (Q \to R)$ is true because P is false, while $P \to Q$ is true and R is false, so the conclusion is false. Chapter 4 proves that such a choice rules out every derivation, so that a search for it can only stall: the chapter's open question, in a new place.

Exercise. R4

Ana and Ben talk about a real number.

> **Ana.** Let x be a real number with $x^2 < x$. Then x is positive.
>
> **Ben.** Why can't it be negative?
>
> **Ana.** A negative number is less than its square.
>
> **Ben.** And zero?
>
> **Ana.** At zero, x^2 equals x. And x is less than 1, because for $x \geq 1$ we would have $x^2 \geq x$.
>
> **Ben.** So $0 < x < 1$. But you never said why $x^2 \geq x$ when $x \geq 1$.
>
> **Ana.** Multiply $x \geq 1$ by x, which is positive.

(a) What is the conclusion, and who states it?
(b) Turn the dialogue into a derivation.
(c) Is Ana's last answer enough?

Solution.

(a) "$0 < x < 1$", stated by Ben. Ana's first line, "x is positive", is an intermediate step.

(b) With, for instance, N: "$x < 0$", Z: "$x = 0$", B: "$x \geq 1$", L: "$x^2 < x$", the premise is L. The facts:

- $N \to \neg L$ (Ana: a negative x has $x^2 > 0 > x$);
- $Z \to \neg L$ (Ana: $x^2 = x$ at 0);
- $B \to \neg L$ (Ana, then justified at Ben's request).

The conclusion $(\neg N \wedge \neg Z) \wedge \neg B$ reads as "$x \geq 0$, $x \neq 0$, $x < 1$" by the convention of Section 2.1, that is, $0 < x < 1$, since $x \geq 0$ and $x \neq 0$ give $x > 0$. The sequent is

$$N \to \neg L,\ Z \to \neg L,\ B \to \neg L,\ L \vdash (\neg N \wedge \neg Z) \wedge \neg B.$$

For the derivation: MT (Proposition 2.3.1) needs the negation of each fact's conclusion, and that is $\neg\neg L$, given by $\neg\neg$I (Proposition 2.5.1 (a)); hence line 5.

1	$N \to \neg L$	
2	$Z \to \neg L$	
3	$B \to \neg L$	
4	L	
5	$\neg\neg L$	$\neg\neg$I 4
6	$\neg N$	MT 1, 5
7	$\neg Z$	MT 2, 5
8	$\neg B$	MT 3, 5
9	$\neg N \wedge \neg Z$	$\wedge$I 6, 7
10	$(\neg N \wedge \neg Z) \wedge \neg B$	$\wedge$I 9, 8

(c) Yes: $x \geq 1$ gives $x > 0$, and multiplying $x \geq 1$ by the positive x gives $x^2 \geq x$.

The method is the extraction of this section (Example (An argument about a prime, extracted)), run on a dialogue; the rules are Chapter 2's. The look-alike, $\neg N$ by $\to$E from line 1 and L, fails: $\to$E needs the hypothesis N on a line, and MT needs $\neg\neg L$, not L, which is why line 5 is written.

Chapter 4

Truth, Soundness and Completeness

Chapter 3 ended on a question: we show that something is derivable by exhibiting a derivation, but how could we ever show that something is **not** derivable? Section 3.1 searched for a derivation of $P \to Q \vdash Q \to P$, and every route stalled. A search that fails proves nothing: a cleverer derivation might exist. Here is something that does prove it.

p	q	$p \to q$	$q \to p$	
T	T	T	T	
T	F	F	T	
F	**T**	**T**	**F**	**this row**
F	F	T	T	

In the third row the premise $p \to q$ (Section 4.1 says why the letters are now lower-case) is true, because its hypothesis is false (the promise reading of Section 1.1), and the conclusion $q \to p$ is false. By the end of this chapter this one row is a **proof** that no derivation of $q \to p$ from $p \to q$ exists, however long.

The answer is to give formulas a **meaning**. **Soundness** says that whatever can be derived is true in every row where the premises are true; so one such row with a false conclusion proves that no derivation exists. **Completeness** says that whatever is true in every such row can be derived. To prove them, formulas become objects and **induction** arrives. Between the two theorems sits a search (Section 4.6) that, for any finite list of premises and any conclusion, finds a refuting row or shows that there is none. The core ends with completeness; normal forms, the constructive content of completeness (second pass), and compactness (optional) follow.

What you need. Chapters 1–3: the rules for every connective, the definition of a derivation with its boxes and its scope rule (Definition 1.4.2), derived rules and Lemma 2.3.3 (Section 2.3), the classical rules of Section 2.5, above all Proposition 2.5.3, the search method of Section 3.1, and the four kinds of task of Section 3.2 with its word *consistent* (Definition 3.2.1). The arithmetic and order of the natural numbers are school facts.

Time. About 45 hours on the core track (Sections 4.1–4.8: all of A, one B and one C per section, the exit checks); about 90 hours for everything.

Minimum before *Sets, Relations, and Cardinality*. Chapters 1–6 and 8, core; Sections

7.1–7.2 before its Chapter 3.

How to read this chapter. A first reading takes Sections 4.1–4.8 with every second-reading fold closed. Each section's main line teaches the section, and every A and B exercise can be done from it. The folds hold the proofs of structural induction, the parenthesis count and structural recursion, the bookkeeping lemmas, the proof that the search of Section 4.6 decides, and the descent in the completeness proof. They are not optional: later proofs use those results, and the proofs are complete. Section 4.9, on normal forms, is a second-pass section; nothing in Chapters 5–8 uses it. Section 4.10, on compactness, is optional; nothing outside it depends on it.

4.0.1 Roadmap

- **Formulas as Objects.** The exact definition of a formula, parse trees, the object language and the metalanguage, and substitution into a derivation.
- **Induction on the Natural Numbers.** The principle, strong induction, a checklist for every induction, and proving more in order to prove less.
- **Induction and Recursion on Formulas.** Structural induction, a count of parentheses, the proof that every formula is read in only one way, and definitions by recursion, which fail without it.
- **Valuations and Truth Tables.** The value of every formula computed from its atoms, 2^n rows for n atoms, truth-functional connectives, and a game for the clauses.
- **Tautologies, Entailment and Equivalence.** $\Gamma \models \psi$, satisfiable and contingent formulas, contradictory and contrary pairs, replacement and substitution, and every question reduced to one: is this list unsatisfiable?
- **Hunting for a Refuting Row.** Forced moves, splits, the search drawn as a tree, and why a closed tree proves an entailment.
- **Soundness.** Every rule checked against the tables, by induction on the lines of a derivation, so that one refuting row rules out every derivation; and the move *test the subgoal*.
- **Completeness.** The rules prove their own truth tables row by row, and excluded middle glues the rows; consistent and satisfiable coincide; a truth-table step, TAUT.
- **Normal Forms and Adequate Connectives** (second pass). Normal forms, adequate connectives, calculating with equivalences, and why a stronger hypothesis gives a weaker theorem.
- **Compactness and Maximal Consistent Sets** (optional). Infinitely many premises, Lindenbaum's construction, and why every finite part satisfiable means the whole is.

4.0.2 Named moves

- **Induction on the shape of a formula.** Prove it for atoms and $\bot$, then for each connective from the parts. (Section 4.3)
- **Check the induction.** Objects, property, base cases, steps, principle. (Section 4.2)
- **Prove more to prove less.** Strengthen the claim until the step goes through. (Sections 4.2, 4.3)
- **Define by recursion on formulas.** One clause per way of building a formula. (Section 4.3)
- **Work backwards from F.** Assume the formula false and follow what is forced. (Section 4.4)
- **Reduce to unsatisfiability.** Every semantic question asks whether some list is unsatisfiable. (Section 4.5)
- **Search for a refuting row.** Forced moves first, split when nothing is forced, close branches early. (Section 4.6)
- **Refute with one row.** One refuting row proves $\Gamma \nvdash \psi$, by soundness. (Section 4.7)

- **Test the subgoal.** Before chasing a subgoal, look for a row that refutes it. (Section 4.7)
- **Induction on the lines of a derivation.** Show that every rule preserves the property. (Section 4.7)
- **Split on an atom.** With $p \vee \neg p$ and $\vee$E, prove the goal under p and under $\neg p$. (Section 4.8)

Where this goes. Chapter 5 looks inside the atoms: it gives "every" and "some" their rules and their meaning. Chapter 8 says what reductio asserts that the other rules do not, and who objected to it. Chapter 7 gives the formulas of predicate logic structures in place of rows; it shows a theory consistent, and an axiom independent, by exhibiting a model (Section 7.4), asks which sets a formula can define (Section 7.5), and states completeness and compactness for predicate logic (Section 7.6, optional). Three later volumes take up threads of this chapter. *Sets, Relations, and Cardinality* makes "the smallest collection of strings closed under the building steps" and "function" exact (Chapter 1). It proves induction equivalent to the well-ordering principle and shows a famous false induction that breaks at its first step (Chapter 2). *Number Systems: From* $\mathbb{N}$ *to* $\mathbb{R}$ derives induction from the axioms of set theory (Chapter 1). *Analysis on the Real Line* takes the least upper bound axiom as given, as this chapter takes induction (Chapter 1).

4.0.3 Summary

Formulas are strings built by the clauses of Definition 4.1.1, and each is read in only one way (Theorem 4.3.3), so a function on formulas can be defined by one clause per connective. Induction, on the numbers and on formulas, is the tool for statements about all of them; its checklist and its strengthening are the two habits it needs. A valuation gives every formula a truth value, and entailment, *satisfiable*, *contingent* and the pair relations are defined from rows. Every semantic question reduces to whether some list is unsatisfiable, and a search tree decides it (Proposition 4.6.1). Soundness (Theorem 4.7.2) says that a refuting row rules out every derivation, and completeness (Theorem 4.8.4) that every valid sequent has a derivation; so $\vdash$ and $\models$ coincide, and so do *consistent* and *satisfiable*, and derivability is decided by a finite table. TAUT records a truth-table step as one line.

All of this is proved for propositional formulas and finitely many premises (infinitely many in the optional Section 4.10). What is open is the inside of an atom: "for every $\varepsilon > 0$ there is an N" is one letter to a truth table, and how to reason inside it is Chapter 5's question. Also still open: which tautologies need reductio (the optional Theorem 4.7.4 settles two), and how short a derivation can be.

4.0.4 Terms

- **adequate set of connectives**: expresses every truth function (Definition 4.9.5).
- **atom**: one letter such as p, q, p_1, a formula alone (Definition 4.1.1).
- **conjunctive, disjunctive normal form**: an "and" of "or"s; an "or" of "and"s (Definition 4.9.3, Definition 4.9.2).
- **contingent formula**: true in some row and false in another (Definition 4.5.1).
- **contradiction**: a formula false in every row (Definition 4.5.1).
- **contradictory, contrary, subcontrary**: exactly one true; never both true; never both false (Definition 4.5.5).
- **entailment**, $\Gamma \models \psi$: no refuting row (Definition 4.5.3).
- **formula**: a string built from atoms and $\bot$ by the connectives (Definition 4.1.1).
- **height**: the nodes on a longest branch of the parse tree (Definition 4.3.1).
- **induction**: from $P(0)$ and each step, $P(n)$ for every n (Axiom 4.2.1).

- **literal**: an atom or the negation of an atom (Definition 4.9.1).
- **parse tree**: the tree showing how a formula is built (Definition 4.1.2).
- **positive, negative occurrence**: a part's sign, flipped by $\neg$ and by hypotheses (Definition 4.9.4).
- **premises, infinitely many**: $\vdash$ and $\models$ for a collection (Definition 4.10.1).
- **refuting row**: premises true and conclusion false (Definition 4.5.3).
- **satisfiable**: some row makes every member true (Definition 4.5.2).
- **search tree**: the hunt for a refuting row, drawn (Definition 4.6.1).
- **semantic equivalence**, $\equiv$: the same value in every row (Definition 4.5.4).
- **subformula**: a formula met while building another (Definition 4.1.3).
- **TAUT**: one line for a truth-table step (Definition 4.8.1).
- **tautology**: a formula true in every row (Definition 4.5.1).
- **truth function**: a truth value for each list of truth values (Definition 4.4.4).
- **truth table**: a formula's value in each of the 2^n rows (Definition 4.4.3).
- **truth value**: T or F, computed from the atoms (Definition 4.4.2).
- **valuation**: a choice of T or F for each atom (Definition 4.4.1).

4.1 Formulas as Objects

Chapter 3 ended on a question that no derivation can answer. We show that a sequent holds by writing a derivation of it; but how could we show that one does **not** hold, that no derivation of $Q \to P$ from $P \to Q$ exists, however long and however clever? This chapter's answer is a theorem about **every** derivation of **every** formula, and such a theorem needs, first of all, an exact description of what a formula is. This section gives it.

4.1.1 Why formulas must become objects

Here is the theorem that Section 4.7 proves, stated loosely:

> If ψ can be derived from premises Γ, then ψ is true in every situation in which all the premises are true.

Read it as a mathematician reads a statement about all integers. "For every integer n, $n^2 - n$ is even" can be proved only by someone who knows what an integer is: every integer is even or odd, and each kind can be handled. A proof needs a way to take hold of an arbitrary object of the kind the statement is about. The loose theorem above is about arbitrary derivations, and a derivation is a list of formulas, so the proof will need a way to take hold of an arbitrary formula.

Chapters 1–3 never needed one. There we read statements by eye: we found the main connective of $\neg(P \wedge Q) \to R$ (Definition 1.1.2) by looking, and we drew its tree by hand. That is fine for one statement at a time. It is useless for a claim about all statements, because "all the strings we would recognize as statements" is not a description anyone can reason from. So we now say exactly which objects we mean, in a form that later proofs can use.

The description has three parts. First, the **alphabet**, the symbols formulas are written with. Second, the **formation rules**, which say how formulas are built. Third, a closing clause saying that **nothing else** is a formula. The third is the one that is easy to forget and that makes everything else work.

4.1.2 The alphabet

Formulas are strings of symbols. A **string** is a finite list of symbols written one after another, and its **length** is the number of symbols in it. The symbols are these.

- **Atoms:** $p, q, r, p_1, p_2, p_3, \ldots$, infinitely many of them. Each atom is **one** symbol: p_1 is a single symbol whose name happens to carry a subscript, not the two symbols p and 1.
- **The constant** $\bot$, read "contradiction", as in Chapter 1.
- **The connectives:** $\neg$ (the only one that takes one part) and $\wedge, \vee, \to, \leftrightarrow$ (the **binary** connectives, which take two).
- **The parentheses** (and).

An atom plays the part that a letter played in Chapters 1–3: a statement whose inner structure we ignore. In those chapters the letters were capitals, P, Q, R, and they stood for **arbitrary** statements, atomic or compound. From now on we keep two kinds of letter apart, because the theorems of this chapter need the difference.

Remark. **Who is who: the letters of this chapter.**

Letter	Stands for	Inside a formula?	Example
$p, q, r, p_1, p_2, \ldots$	an **atom**, one symbol of the language	yes	$(p \wedge \neg q)$
φ, ψ, χ (phi, psi, chi)	any formula: a **metavariable**, a name we use when talking about formulas	never	"if φ is a formula, so is $\neg\varphi$"
Γ, Δ (capital gamma, delta)	a finite list of formulas, as in the sequents of Chapter 1 (Definition 1.2.3)	never	$\Gamma \vdash \psi$
P, Q, R of Chapters 1–3	arbitrary statements; read from now on as atoms, or as any formulas (the last part of this section)	they are replaced	$P \to Q$ read as $p \to q$

Chapter 1 already used the three letters φ, ψ, χ in exactly this way, in the schemas of its rules (Definition 1.2.1).

Quick check. Which of these is a formula, and which is a name for formulas or for a list: $(p \to q)$; $(\varphi \to q)$; Γ?

Solution. $(p \to q)$ is a formula. $(\varphi \to q)$ is not a formula, since φ is no symbol of the language; it is a pattern, and each formula put in for φ gives a formula. Γ names a list of

formulas.

Two languages. Everything in this chapter is said in two languages at once. The **object language** is the language of formulas: its expressions are strings over the alphabet above, such as $(p \to q)$. The **metalanguage** is the language in which we talk about those strings: English with ordinary mathematics, together with names for formulas (φ, ψ) and for lists (Γ), and claims such as $\Gamma \vdash \psi$ and, from Section 4.5 on, $\Gamma \models \psi$. Neither $\vdash$ nor $\models$ is a symbol of the alphabet. The sentence "$(p \to q)$ has length 5" belongs to the metalanguage and speaks about a string of the object language; $(p \to q)$ on its own is a string, not a claim. The theorems of this chapter, unique readability, soundness and completeness, are theorems **about** the object language, proved in the metalanguage by ordinary mathematics: induction, cases, counting. To speak about a symbol we name it, as we name a word. "p is an atom" mentions the symbol p, in the way that " 'seven' has five letters" mentions a word.

Quick check. Sort into strings of the object language, claims of the metalanguage, and neither:

(a) $((p \wedge q) \to p)$
(b) $p \wedge q \vdash p$
(c) the main connective of $((p \wedge q) \to p)$ is $\to$
(d) $(\varphi \wedge q)$
(e) $\neg\bot$ is a formula

Solution. (a) A string of the object language: a formula.

(b) A claim of the metalanguage, since $\vdash$ is not a symbol of the alphabet. It is true, by $\wedge$E.

(c) A claim of the metalanguage, and true.

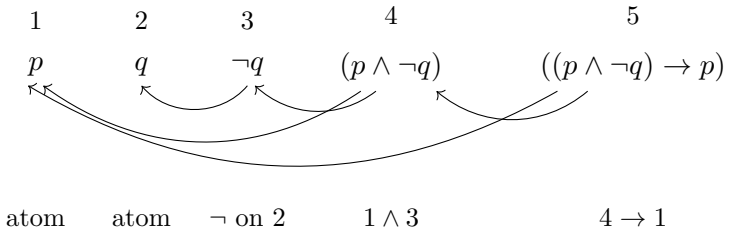
(d) Neither. The letter φ is a metavariable, not a symbol, so the expression is a pattern; each formula put in for φ gives a formula.

(e) A claim of the metalanguage, and true: $\bot$ is a formula, and $\neg$ in front of a formula gives a formula (clauses (F1) and (F2) of the definition below).

4.1.3 Formulas

Chapter 1 described a statement as something "built from smaller ones by a connective". The definition below says exactly how the building goes. It is an **inductive definition**: it names the starting pieces, names the ways of building new formulas from old, and then closes the door on everything else.

Building one formula. Here is the formula $((p \wedge \neg q) \to p)$ built one step at a time.



Each entry is an atom, or is built from earlier entries by one connective. The atom p is used twice and listed once.

A list like this is called a formation sequence. The definition below has one more clause than the building steps. Its clause (F4) says that the formulas are exactly the last entries of such lists: the list you just saw, in general.

In clause (F4) below, $\circ$ stands for any one of the four binary connectives, as φ stands for any formula: $(\chi_j \circ \chi_l)$ means any one of $(\chi_j \wedge \chi_l)$, $(\chi_j \vee \chi_l)$, $(\chi_j \to \chi_l)$, $(\chi_j \leftrightarrow \chi_l)$.

A formula is a string built from atoms and $\bot$ by the connectives, one step at a time, with a pair of parentheses around every binary step.

Definition 4.1.1. Formula

The **formulas** are the strings given by the following clauses.

(F1) Every atom is a formula, and $\bot$ is a formula.

(F2) If φ is a formula, then $\neg\varphi$ is a formula.

(F3) If φ and ψ are formulas, then $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \to \psi)$ and $(\varphi \leftrightarrow \psi)$ are formulas.

(F4) **Nothing else** is a formula: a string is a formula **only if** it is the last entry of a **formation sequence**. A formation sequence is a finite list of strings $\chi_1, \ldots, \chi_m$, with $m \geq 1$, in which **every** entry χ_i is an atom or $\bot$, or is $\neg\chi_j$ for some **earlier** entry χ_j (so $j < i$), or is $(\chi_j \circ \chi_l)$ for some earlier entries χ_j and χ_l (so $j, l < i$; $j = l$ is allowed) and some binary connective $\circ$.

In words, clause by clause.

- *(F1), the starting pieces.* A single atom is a formula, and so is $\bot$. These are the only formulas of length one.
- *(F2), negation.* Writing $\neg$ in front of a formula gives a formula. No parentheses are added: $\neg p$ and $\neg\neg p$ are formulas, and $\neg\neg p$ is $\neg$ in front of the formula $\neg p$.
- *(F3), the binary connectives.* Writing a binary connective between two formulas **and** wrapping the result in one pair of parentheses gives a formula. The two formulas may be the same: $(p \wedge p)$ is a formula.
- *(F4), nothing else.* Clauses (F1)–(F3) say which strings **are** formulas. They do not say that other strings are **not**. Clause (F4) does: a formula is exactly something we can reach from atoms and $\bot$ by finitely many uses of (F2) and (F3), and the formation sequence is the written record of such a reaching. Think of building blocks: (F1) lists the blocks, (F2) and (F3) the two ways of joining them, and (F4) says that only what is built in this way counts.

Two facts about formation sequences are used constantly, and both are immediate from the definition.

First, **every entry of a formation sequence is a formula.** Walk along the list. The first entry χ_1 is an atom or $\bot$, so it is a formula by (F1). Each later entry is an atom or $\bot$, a formula by (F1), or it is $\neg\chi_j$ or $(\chi_j \circ \chi_l)$ for earlier entries that are already formulas, and so it is a formula by (F2) or (F3). (Section 4.3 makes this walk exact with strong induction, in the fold "Sets built from a base".)

Second, **every formula was made by one of the clauses from formulas.** Take a formula χ and a formation sequence ending in it. Its last entry χ is an atom or $\bot$, or it is $\neg\chi_j$, or it is $(\chi_j \circ \chi_l)$, where the earlier entries χ_j, χ_l are formulas by the first fact. So every formula is an atom, or $\bot$, or $\neg\varphi$ for a formula φ, or $(\varphi \circ \psi)$ for formulas φ, ψ. We will call this **looking at the last step** of a formula.

Looking at the last step already tells us something about the first symbol. A formula of the form $\neg\varphi$ begins with $\neg$, and one of the form $(\varphi \circ \psi)$ begins with (and ends with). So:

> **The first-symbol test.** A formula of length one is an atom or $\bot$. A longer formula begins with $\neg$ or with (, and if it begins with (, it is $(\varphi \circ \psi)$ for formulas φ, ψ and ends with).

(A formula of the form $\neg\varphi$ has length at least 2, and one of the form $(\varphi \circ \psi)$ has length at least 5, since φ and ψ each have at least one symbol. So a formula of length one is made by (F1).)

Quick check. Write a formation sequence for $\neg(p \vee \bot)$, and say which earlier entries each entry uses.

Solution. p, $\bot$, $(p \vee \bot)$, $\neg(p \vee \bot)$. Entry 3 uses entries 1 and 2, and entry 4 uses entry 3; entries 1 and 2 use none.

Example. Formation sequences

Write down a formation sequence for each of the strings p, $\bot$, $\neg\neg q$ and $(\neg p \vee (q \to \bot))$, which shows that each is a formula.

Solution. For p, the one-entry list p is a formation sequence: its only entry is an atom. Likewise the list $\bot$. These are the degenerate cases, and they matter: a formula need not have been built at all. Every structural argument in Section 4.3 has to deal with them.

For $\neg\neg q$:

$$q, \quad \neg q, \quad \neg\neg q.$$

The first entry is an atom; the second is $\neg$ in front of the first; the third is $\neg$ in front of the second.

For $(\neg p \vee (q \to \bot))$:

$$p, \quad \neg p, \quad q, \quad \bot, \quad (q \to \bot), \quad (\neg p \vee (q \to \bot)).$$

Entries 1, 3 and 4 are atoms or $\bot$. Entry 2 is $\neg$ in front of entry 1. Entry 5 is $(\chi_3 \to \chi_4)$, and entry 6 is $(\chi_2 \vee \chi_5)$.

Second reading · *Many formation sequences*

A formula has many formation sequences. In Example (Formation sequences) we could have listed q before p in the last list, or added entries that are never used, such as r or $\neg r$. The definition asks only that **some** formation sequence ends in the string, and nothing in this chapter depends on which one we take.

To show that a string **is** a formula, we exhibit a formation sequence, just as we show that a sequent holds by exhibiting a derivation. To show that a string is **not** a formula, exhibiting is not enough; we need a property that every formula has and the string lacks. The first-symbol test supplies such properties.

Example. Strings that are not formulas

Show that none of the strings $p\wedge \to q$, $(p \wedge q$, $(\neg p)$ and $p \wedge q$ is a formula.

Solution. $p\wedge \to q$: it has length 4, so by the first-symbol test a formula would begin with $\neg$ or (. It begins with p. Not a formula.

$(p \wedge q$: it begins with (, so by the first-symbol test a formula would end with). It ends with q. Not a formula. (The missing parenthesis is the whole problem: $(p \wedge q)$ is a formula.)

$(\neg p)$: it begins with (, so a formula would have to be $(\varphi \circ \psi)$, with the two symbols $\neg p$ between the parentheses equal to $\varphi \circ \psi$, and so containing a binary connective. The string $\neg p$ contains none. Not a formula. The official language puts no parentheses around a negation; (F2) adds none.

$p \wedge q$: it has length 3 and begins with p. Not a formula. This one is a surprise at first, since we have written $p \wedge q$ all along. It is **shorthand** for the formula $(p \wedge q)$, in the way explained later in this section.

The last two strings are ones Chapter 1 wrote freely, and that is no mistake in Chapter 1. Chapter 1 used a convenient shorthand; the official formulas carry every parenthesis that (F3) puts in and none that it does not. We come back to the shorthand below.

Quick check. Which of the following strings are formulas? For each, give a formation sequence or name the test it fails.

(a) $\neg(p \to \bot)$
(b) $(p \to \neg)$
(c) $\bot\bot$

Solution. *(a)* A formula: p, $\bot$, $(p \to \bot)$, $\neg(p \to \bot)$ is a formation sequence.

(b) Not a formula. It begins with (, so it would be $(\varphi \to \psi)$ with φ the string p and ψ the

string ¬ (the only binary connective is the →). But ¬ is a string of length one that is neither an atom nor ⊥, so by the first-symbol test it is not a formula.

(c) Not a formula. It has length 2, so it would begin with ¬ or (; it begins with ⊥.

4.1.4 Why this definition

Each feature of the definition is there because something breaks without it.

Without (F4), clauses (F1)–(F3) would be satisfied by far too many collections of strings. The collection of **all** strings contains every atom and ⊥, and is closed under writing ¬ in front and under wrapping two strings with a connective in parentheses; so (F1)–(F3) alone cannot tell us that $p\wedge \to q$ is excluded. Clause (F4) says the formulas are only those strings that the building steps reach. In other words, the formulas are the **smallest** collection of strings closed under (F1)–(F3).

Without the parentheses in (F3), a string could be read in two ways. Take $p \wedge q \vee r$. Built one way, it is "$p \wedge q$" joined to r by $\vee$; built the other way, it is p joined to "$q \vee r$" by $\wedge$. The two constructions have different parse trees (defined below), and they say different things (the everyday reading of Chapter 1 is enough here): if p is false and r is true, then "p and q, or r" is true, because r is, while "p, and q or r" is false, because p is. With the parentheses of (F3) the two constructions give two **different** strings,

$$((p \wedge q) \vee r) \qquad \text{and} \qquad (p \wedge (q \vee r)),$$

and the ambiguity disappears. Section 4.3 proves that it disappears for every formula: a formula can be built in only one way (the unique readability theorem).

↻ **Second reading** · *Two smaller points*

Why not parentheses around a negation too? Nothing would break, but nothing would be gained. The symbol ¬ always stands in front of what it negates, so $\neg\varphi$ can be read in only one way without help. Keeping the official language as lean as possible makes the proofs about it shorter.

Why atoms, and not sentences like "$x > 0$"? Because the theorems of this chapter are about **forms**, and the form of an argument is what remains when the content of its atomic statements is forgotten (Chapter 1's slogan: logic studies the forms of arguments). An atom is a statement reduced to a name.

Warning. A formula is a string of symbols, and questions about formulas are questions about strings. "Is $(p\wedge p)$ the same formula as p?" has the answer **no**: the first has length 5 and the second length 1, even though they always have the same truth value (Section 4.4 gives the means to check this). Sameness of formulas is letter-for-letter identity; sameness of meaning is a different relation, which Section 4.5 names.

4.1.5 Parse trees

A formation sequence records how a formula was built, but it is a list, and a list hides the shape of the building. Chapter 1 drew the shape as a tree, with the main connective at the top (the tree after Definition 1.1.2). We now make that drawing official.

A parse tree is a formula's construction drawn upside down: the last step at the root, the atoms at the leaves.

Definition 4.1.2. Parse Tree

A **parse tree for** a formula is defined along the clauses of Definition 4.1.1.

(T1) For an atom or $\bot$: a single node, labeled by that atom or by $\bot$.

(T2) If T is a parse tree for φ, then the tree whose **root** is labeled $\neg$ and has T hanging below it as its only subtree is a parse tree for $\neg\varphi$.

(T3) If T and T' are parse trees for φ and ψ, then the tree whose root is labeled $\circ$, with T hanging below it on the **left** and T' on the **right**, is a parse tree for $(\varphi \circ \psi)$.

(T4) **Nothing else** is a parse tree: a tree is a parse tree for a formula φ **only if** that follows from finitely many uses of (T1)–(T3). In particular, a parse tree for $\neg\varphi$ is made by (T2), and a parse tree for $(\varphi \circ \psi)$ by (T3).

The **leaves** are the nodes with nothing below them; a **branch** is a path from the root down to a leaf.

In words: (T1) is the tree of a formula that was not built at all, a single node. (T2) puts a new root $\neg$ on top of an existing tree. (T3) joins two existing trees under a new root carrying the connective, keeping left and right apart, because $(p \to q)$ and $(q \to p)$ are different formulas. (T4), as (F4) did for formulas, says that only trees built this way count; Section 4.3 uses it when it asks how a given tree could have been made. Parentheses do not appear in a parse tree: their only job in the string was to show the grouping, and the tree shows the grouping by its shape.

Every formula has a parse tree. Take a formation sequence of the formula and build a tree for each entry in turn: an atom or $\bot$ gets (T1), and an entry $\neg\chi_j$ or $(\chi_j \circ \chi_l)$ gets (T2) or (T3) from the trees already built for the earlier entries. The tree built for the last entry is a parse tree for the formula.

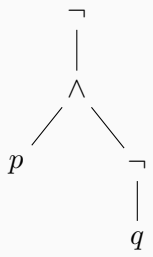
Could a formula have **two** different parse trees? That would happen if one string could be built in two different ways, as $p \wedge q \vee r$ could if parentheses were not written. Section 4.3 proves that it never happens (the unique readability theorem); until then we say "**a** parse tree". After that section, every formula has exactly one, and we speak of **the** parse tree.

Example. Three parse trees

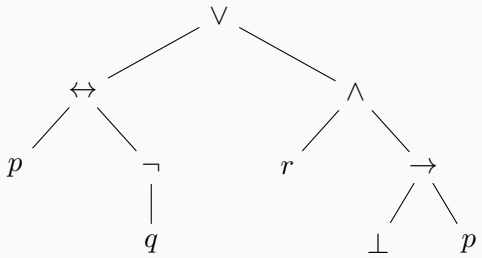
Draw a parse tree for each formula: (a) p; (b) $\neg(p \wedge \neg q)$; (c) $((p \leftrightarrow \neg q) \vee (r \wedge (\bot \to p)))$, which uses every connective and $\bot$.

Solution. (a) A single node labeled p, by (T1). This is the degenerate tree: its root is also its only leaf, and its only branch has one node.

(b) By (T1) the atoms p and q get one-node trees; (T2) puts $\neg$ above q; (T3) joins p and $\neg q$ under $\wedge$; (T2) puts $\neg$ on top.



(c) The main connective is the $\vee$: it is the last step, joining $(p \leftrightarrow \neg q)$ on the left to $(r \wedge (\bot \to p))$ on the right. Each side is taken apart in the same way.



The leaves are p, q, r, $\bot$ and p again: the atom p occurs twice in the formula, and each occurrence is its own leaf.

To draw a parse tree from a string, find the last step, draw it as the root, and repeat on the parts. For a formula $(\varphi \circ \psi)$, the last step is the binary connective that sits inside **exactly one** pair of parentheses; in (c), it is the $\vee$, and every other binary connective sits inside at least two pairs. Section 4.3 proves that this recipe always finds the main connective, and only it.

4.1.6 Subformulas

Each node of a parse tree is the root of the part of the tree below it, and that part is itself a parse tree, built by the same clauses, for a formula. The formulas that arise in this way are the pieces a formula is built from.

Definition 4.1.3. Subformula

The **subformulas** of a formula φ are the formulas whose parse trees hang from the nodes of a parse tree for φ: for each node, the formula of which the part of the tree from that node downwards is a parse tree. The formulas at the children of the root are the **immediate subformulas** of φ.

Until Section 4.3 this depends, as far as we know, on the parse tree chosen; Section 4.3 shows there is only one, so the subformulas of φ are determined by φ.

So the subformulas of φ include φ itself (at the root) and each atom that occurs in φ (at the leaves). A formula of the form $\neg\varphi$ has one immediate subformula, φ; a formula $(\varphi \circ \psi)$ has the two immediate subformulas φ and ψ (one, if φ and ψ are the same formula); an atom or $\bot$ has none. When we list the subformulas, a formula that sits at several nodes is listed once.

For the formula (c) of Example (Three parse trees), the subformulas are

$$((p \leftrightarrow \neg q) \vee (r \wedge (\bot \to p))), \quad (p \leftrightarrow \neg q), \quad (r \wedge (\bot \to p)),$$
$$\neg q, \quad (\bot \to p), \quad p, \quad q, \quad r, \quad \bot,$$

nine in all. The tree has ten nodes, $\vee$, $\leftrightarrow$, p, $\neg$, q, $\wedge$, r, $\rightarrow$, $\bot$ and p again, and only p sits at two of them; so ten nodes give nine distinct subformulas. Counting by eye is easy to get wrong, and the safe method is to walk the tree node by node, as we just did.

A string that merely **appears inside** a formula need not be a subformula. The formula $((p \wedge q) \vee r)$ contains the string $q) \vee r$, which is not a formula at all, and it contains the string $p \wedge q$, which is only the shorthand of a formula. Its subformulas are the formulas at the nodes: $((p \wedge q) \vee r)$, $(p \wedge q)$, p, q and r.

Quick check. Draw the parse tree of $(\neg p \vee q) \leftrightarrow \neg(r \wedge p)$ (in official form, $((\neg p \vee q) \leftrightarrow \neg(r \wedge p))$) and list its subformulas.

Solution. The root is $\leftrightarrow$. Its left subtree is $\vee$ over $\neg p$ and q, where $\neg p$ is $\neg$ over the leaf p. Its right subtree is $\neg$ over $\wedge$, and $\wedge$ has the leaves r and p.

The subformulas are the formula itself, $(\neg p \vee q)$, $\neg(r \wedge p)$, $\neg p$, q, $(r \wedge p)$, p and r: **eight**, with p listed once although it sits at two leaves.

Warning. A parse tree is not a truth table. The tree records how a formula is **built**, and it is the same whatever the atoms stand for; nothing in it is true or false. Values come in Section 4.4, where they are computed **along** the tree, from the leaves up. Mixing the two, for instance writing "T" beside a node while drawing the tree, is a habit worth not forming: the same tree carries different values under different assignments.

4.1.7 Shorthand and official formulas

Nobody wants to write $(((p \wedge q) \rightarrow r) \leftrightarrow (p \rightarrow (q \rightarrow r)))$. Chapter 1 used conventions that let us drop parentheses (Section 1.1, the conventions before Definition 1.1.2), and we keep using them. What changes is their status: they are **not** part of the definition of a formula. A string written with the conventions is **shorthand**, and it stands for exactly one official formula, the one obtained by putting the dropped parentheses back.

The translation, from shorthand to official form, goes as follows.

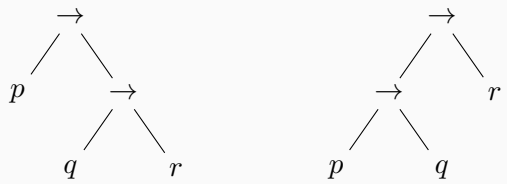
1. Put back the outer pair of parentheses of a formula whose main connective is binary: $p \wedge q$ stands for $(p \wedge q)$.
2. $\neg$ binds most tightly: it applies to the smallest formula that follows it. So $\neg p \wedge q$ stands for $(\neg p \wedge q)$, not for $\neg(p \wedge q)$.
3. $\wedge$ and $\vee$ bind more tightly than $\rightarrow$ and $\leftrightarrow$: $p \wedge q \rightarrow r$ stands for $((p \wedge q) \rightarrow r)$.
4. $\rightarrow$ groups to the right: $p \rightarrow q \rightarrow r$ stands for $(p \rightarrow (q \rightarrow r))$.
5. Chapter 1 sometimes wrote a negation in parentheses for emphasis, as in $(\neg P) \rightarrow Q$; such parentheses are simply removed.

In a mixture of $\wedge$ with $\vee$, or of $\rightarrow$ with $\leftrightarrow$, the shorthand keeps the parentheses (Chapter 1's rule for such mixtures), so no rule is needed for those.

Example. Two groupings of three implications

Write the official formula for the shorthand $p \to q \to r$, draw its parse tree beside that of $((p \to q) \to r)$, and find a situation in which the two formulas say different things.

Solution. By rule 4, $p \to q \to r$ stands for $(p \to (q \to r))$: the first $\to$ is the main connective, and its conclusion is $(q \to r)$. The other grouping, $((p \to q) \to r)$, has the second $\to$ as main connective, and its hypothesis is $(p \to q)$. Left, $(p \to (q \to r))$; right, $((p \to q) \to r)$.



The two trees are mirror images in shape, and they say different things. Suppose p, q and r are all false. The left formula is an implication whose hypothesis p is false, so it is true (vacuous truth, Section 2.1). The right formula has the true hypothesis $p \to q$ (again vacuous, since p is false) and the false conclusion r, so it is false. The everyday reading of Chapters 1 and 2 is enough to see this; Section 4.4 makes it a computation.

This is why the grouping convention matters. Grouping to the right is the useful choice, because statements of the form "if P, then if Q, then R" are common and $(P \to Q) \to R$ is rare.

Going the other way, from official to shorthand, we drop every pair of parentheses the rules would put back. $(\neg(p \vee q) \leftrightarrow (\neg p \wedge \neg q))$ becomes $\neg(p \vee q) \leftrightarrow \neg p \wedge \neg q$: the outer pair goes by rule 1, the pair around $\neg p \wedge \neg q$ by rule 3, and the pair after the first $\neg$ must stay, since without it rule 2 would read $\neg p \vee q$.

From now on we write formulas in shorthand whenever that is clearer, and every statement about "the formula $p \wedge q$" is about the official formula $(p \wedge q)$. Where the official form matters, as in counting parentheses or symbols, we write it out.

4.1.8 Letters, metavariables and substitution

We can now say exactly what the letters of Chapters 1–3 were.

Every rule was stated as a schema (Definition 1.2.1), with metavariables: $\wedge$I writes $\varphi \wedge \psi$ from φ and ψ, **for any** formulas φ and ψ. The derivations of those chapters were written with capital letters, as in the derivation of Proposition 1.4.1:

1	$P \to Q$	
2	$Q \to R$	
3	$\quad P$	asm, for $\to$I
4	$\quad Q$	$\to$E 1, 3
5	$\quad R$	$\to$E 2, 4
6	$P \to R$	$\to$I 3–5

In the language of this chapter, read the letters P, Q, R as three atoms. Then this is a derivation

of $p \to q,\ q \to r \vdash p \to r$ in the official sense, written in shorthand. But Chapter 1 meant more: the derivation was a **pattern**, good for any statements put in for the letters. Here is that promise made exact.

A derivation written with letters is a template.

Substitution. Replace each atom of a derivation by a formula, **the same formula at every occurrence of the same atom**, and all such replacements at once. The result is again a derivation.

The fold below checks that every line stays a formula and stays justified.

↻ **Second reading** · *Why the result is again a derivation*

Here is why. First, each line becomes a formula. Replacing atoms by formulas in a formula χ gives a formula: take a formation sequence of χ, write in front of it a formation sequence of each formula put in for an atom, and replace the atoms inside the old entries. Each old entry that was an atom has become the last entry of one of the sequences written in front; each old entry built by (F2) or (F3) is now built by the same clause from the replaced versions of the same earlier entries. So the new list is a formation sequence, and its last entry is the replaced χ.

Second, each line is still justified. Each line of the old derivation was a premise, an assumption, a repetition, or an instance of a rule's schema (Definition 1.4.2, clauses (D1)–(D4)). Being an instance means: the line and the lines it cites are the schema with its metavariables replaced by certain formulas. The schema's own symbols are connectives, $\bot$ and parentheses, never atoms, so replacing atoms in the lines is the same as replacing atoms inside those certain formulas. The new lines are therefore the same schema with its metavariables replaced by the new formulas: an instance of the same rule.

Third, nothing about citations changes. The line numbers, the boxes, and which line cites which are untouched, and whether a line or a box is available depends on nothing else (Definition 1.4.2, the scope rule). So the result is a derivation, with the replaced premises and the replaced last line.

For instance, put $(p_1 \wedge p_2)$ for P, $\neg p_1$ for Q, and $(p_2 \vee p_1)$ for R in the derivation above:

1	$(p_1 \wedge p_2) \to \neg p_1$	
2	$\neg p_1 \to (p_2 \vee p_1)$	
3	$\quad p_1 \wedge p_2$	asm, for $\to$I
4	$\quad \neg p_1$	$\to$E 1, 3
5	$\quad p_2 \vee p_1$	$\to$E 2, 4
6	$(p_1 \wedge p_2) \to (p_2 \vee p_1)$	$\to$I 3–5

Line 4 is still $\to$E: lines 1 and 3 are $\varphi \to \psi$ and φ with φ the formula $p_1 \wedge p_2$ and ψ the formula $\neg p_1$. Line 6 is still $\to$I with the same box.

The fact is immediate rule by rule, because every rule was stated for arbitrary formulas from the start. It is what the phrase "the letters stand for arbitrary statements" (Chapter 1, Sec-

tion 1.2) and the clause "replace each letter ... the same statement for every occurrence" of Definition 2.3.1 were relying on. We use it twice later in the chapter, in the soundness proof of Section 4.7 and in the completeness proof of Section 4.8.

Remark. The replacement must be **uniform**. Replacing only the first P of the derivation above by $p_1 \wedge p_2$, and leaving the P on line 3 alone, would break line 4, which needs the hypothesis of line 1 to be literally the formula on line 3.

Remark (Looking ahead). Section 4.3 proves that every formula is built in only one way. Then "the parse tree" of a formula makes sense, and with it the height of a formula.

4.1.9 Exercises

4.1.9.1 A. Check your understanding

Exercise. A1

(a) Decide which of the following strings are formulas. Give a formation sequence for each formula, and for each nonformula name the test or clause it fails: (i) $(p \vee \bot)$; (ii) $\neg\neg p$; (iii) $(p\neg q)$; (iv) $((p \to q))$; (v) $(p \wedge q \wedge r)$.

(b) List the subformulas of $(\neg(p \to q) \vee q)$.

(c) Write a formation sequence of $((p \vee q) \wedge \neg(p \vee q))$ in which $(p \vee q)$ is listed once.

(d) Why does clause (F3) of Definition 4.1.1 put parentheses around its result, while (F2) does not?

(e) Explain why clauses (F1)–(F3) alone do not show that $\to p$ is not a formula, and which clause does.

Solution. (a) (i) A formula: p, $\bot$, $(p \vee \bot)$.

(ii) A formula: p, $\neg p$, $\neg\neg p$.

(iii) Not a formula. It begins with (, so by the first-symbol test it would be $(\varphi \circ \psi)$, and the string $p\neg q$ between the parentheses would contain a binary connective. It contains none.

(iv) Not a formula. It begins with (, so it would be $(\varphi \circ \psi)$ with $\varphi \circ \psi$ the string $(p \to q)$. The only binary connective in that string is the $\to$, so φ would be the string $(p$, of length 2. A formula of length 2 begins with $\neg$, by the first-symbol test (one beginning with (has length at least 5). So $(p$ is not a formula, and neither is the string. So here the doubled outer pair is not official.

(v) Not a formula. It would be $(\varphi \wedge \psi)$ with the $\wedge$ one of the two. If it is the first, ψ is the string $q \wedge r$; if the second, φ is $p \wedge q$. Each has length 3 and begins with an atom, so by the first-symbol test neither is a formula. (Chapter 1's conventions give no reading of the shorthand $p \wedge q \wedge r$; Section 4.5 introduces it for the right-grouped $(p \wedge (q \wedge r))$ and explains why either grouping would do.)

(b) The formula is $(\neg(p \to q) \vee q)$. Its subformulas are $(\neg(p \to q) \vee q)$, $\neg(p \to q)$, $(p \to q)$, p and q: five, with q listed once although it sits at two leaves.

(c) p, q, $(p \vee q)$, $\neg(p \vee q)$, $((p \vee q) \wedge \neg(p \vee q))$. Entry 4 uses entry 3, and entry 5 uses entries 3 and 4: $(p \vee q)$ is built once and used twice.

(d) A binary connective joins two formulas, and without parentheses a string such as $p \wedge q \vee r$ could be built in two ways, with different meanings (see "Why this definition"). A negation stands in front of the one formula it negates, so there is nothing to group.

(e) Clauses (F1)–(F3) say which strings are formulas and never which are not: the collection of all strings satisfies them. Clause (F4) says a formula is the last entry of a formation sequence, and from it follows the first-symbol test: $\rightarrow p$ has length 2 and does not begin with $\neg$ or (.

Exercise. A2: Drill

Answer each part in a line or two.

(a) Is $\neg\neg\neg p$ a formula?

(b) Is $(p \wedge (q))$ a formula?

(c) How many symbols has $((p \rightarrow q) \vee \neg r)$?

(d) Is $p \wedge q \vdash p$ a formula?

(e) What is the main connective of $\neg(p \wedge q) \rightarrow r$?

(f) Which of these are subformulas of $((p \wedge q) \rightarrow r)$: $(p \wedge q)$; the string $q) \rightarrow r$; $(q \rightarrow r)$; r?

(g) Add one symbol to $(p \wedge \neg q$ to make a formula.

(h) Delete as few symbols as possible from $((\neg p))$ to make a formula.

(i) Insert one pair of parentheses into $(p \vee q \vee r)$ to make a formula, in two different ways.

(j) Write the official formula for $\neg p \vee q \rightarrow r$.

(k) Draw the parse tree of $(p \rightarrow (q \rightarrow p))$ and list its leaves.

(l) In the derivation of Proposition 1.4.1 printed in this section, put q for P, q for Q and p for R. Is the result a derivation?

Solution.

(a) Yes: p, $\neg p$, $\neg\neg p$, $\neg\neg\neg p$ is a formation sequence.

(b) No. It would be $(\varphi \wedge \psi)$ with ψ the string (q). That string begins with (and has length 3, and a formula beginning with (has length at least 5.

(c) Ten: ((p $\rightarrow$ q) $\vee$ $\neg$ r).

(d) No. It is a claim of the metalanguage about two formulas; $\vdash$ is not a symbol of the alphabet.

(e) $\rightarrow$; officially the formula is $(\neg(p \wedge q) \rightarrow r)$.

(f) $(p \wedge q)$ and r. The string $q) \rightarrow r$ is not a formula. $(q \rightarrow r)$ is a formula, but no node of the tree carries it.

(g) $(p \wedge \neg q)$: the string begins with (, so by the first-symbol test it must end with).

(h) Delete all four parentheses, leaving $\neg p$. The string has no binary connective, and a pair of parentheses is official only around a binary step; so the only formulas made of its symbols are $\neg p$ and p.

(i) $((p \vee q) \vee r)$ and $(p \vee (q \vee r))$.

(j) $((\neg p \vee q) \to r)$.

(k) Root $\to$; on the left the leaf p, on the right $\to$ over q and p. The leaves are p, q, p.

(l) Yes. Lines 1–6 become $q \to q$, $q \to p$, q, q, p, $q \to p$, and lines 4, 5, 6 are still instances of $\to$E, $\to$E, $\to$I. A substitution may put the same formula for two letters.

4.1.9.2 B. Practice

Exercise. B1: Trees from Chapter 1

Each formula below appears, with capital letters, in a result of Chapter 1 (Proposition 1.4.2 and Proposition 1.6.1); here the letters are atoms. For each, write the official formula, draw its parse tree, and list its subformulas.

(a) $(p \wedge q) \to r$

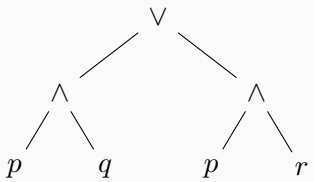
(b) $p \to (q \to r)$

(c) $(p \wedge q) \vee (p \wedge r)$

Solution. *(a)* Official form $((p \wedge q) \to r)$. Root $\to$; left subtree $\wedge$ with leaves p, q; right subtree the leaf r. Subformulas: $((p \wedge q) \to r)$, $(p \wedge q)$, p, q, r.

(b) Official form $(p \to (q \to r))$. Root $\to$; left the leaf p; right subtree $\to$ with leaves q, r. Subformulas: $(p \to (q \to r))$, $(q \to r)$, p, q, r. (This is the left tree of Example (Two groupings of three implications). The formulas of (a) and (b) have different trees; Proposition 1.4.2 says that each can be derived from the other.)

(c) Official form $((p \wedge q) \vee (p \wedge r))$. Root $\vee$; left subtree $\wedge$ with leaves p, q; right subtree $\wedge$ with leaves p, r.



Subformulas: $((p \wedge q) \vee (p \wedge r))$, $(p \wedge q)$, $(p \wedge r)$, p, q, r: six, with p listed once although it sits at two leaves.

Exercise. B2: Shorthand and back

(a) Write in shorthand, dropping every pair of parentheses the conventions allow: $((p \wedge q) \to (r \to p))$ and $(\neg(p \vee q) \to (\neg p \wedge \neg q))$.

(b) Write the official formula for: $p \to q \vee r \to \neg p$ and $\neg p \wedge q \leftrightarrow r$.

Solution. *(a)* $p \wedge q \to r \to p$. The outer pair goes by rule 1; $(p \wedge q)$ loses its pair because $\wedge$ binds more tightly than $\to$ (rule 3); $(r \to p)$ loses its pair because $\to$ groups to the right (rule 4).

$\neg(p \vee q) \to \neg p \wedge \neg q$. The outer pair goes; the pair around $\neg p \wedge \neg q$ goes by rule 3. The pair

after the first ¬ stays: without it, rule 2 would make the ¬ apply to p alone.

(b) $(p \to ((q \vee r) \to \neg p))$. Rule 3 groups $q \vee r$ first; rule 4 then makes the first $\to$ the main connective, with conclusion $((q \vee r) \to \neg p)$.

$((\neg p \wedge q) \leftrightarrow r)$. Rule 2 makes ¬ apply to p alone; rule 3 groups $\neg p \wedge q$ before $\leftrightarrow$.

4.1.9.3 C. Going deeper

Exercise. C1: Forgetting the grouping convention

Suppose we forgot rule 4 of the shorthand, so that $p \to q \to r \to s$ could stand for any official formula obtained by inserting parentheses into it.

(a) List every official formula it could stand for, and draw the shape of each parse tree.
(b) How many are there? Say how the count is found, by splitting at the main connective.
(c) Which of them is the one rule 4 chooses?

Solution (Hint 1). The main connective is one of the three arrows. Choose it first.

Solution (Hint 2). Write c_n for the number of groupings of a chain of n atoms, and find c_1, c_2 and c_3 first.

Solution. (a) The main connective is one of the three arrows; what lies on each side is then grouped in every possible way.

- Main connective the **first** arrow: p on the left, and $q \to r \to s$ on the right, grouped in one of two ways. This gives $(p \to (q \to (r \to s)))$ and $(p \to ((q \to r) \to s))$.
- Main connective the **second** arrow: $(p \to q)$ on the left and $(r \to s)$ on the right, each grouped in one way. This gives $((p \to q) \to (r \to s))$.
- Main connective the **third** arrow: $p \to q \to r$ on the left, grouped in one of two ways, and s on the right. This gives $((p \to (q \to r)) \to s)$ and $(((p \to q) \to r) \to s)$.

Their parse trees have the root $\to$ in each case: in the first two, the leaf p on the left and a chain of two arrows on the right (leaning right, or leaning left); in the third, two one-arrow trees side by side; in the last two, a chain of two arrows on the left and the leaf s on the right.

(b) There are 5. Write c_n for the number of groupings of a chain of n atoms joined by one binary connective. A chain of one or two atoms has one grouping: $c_1 = c_2 = 1$. For three atoms the main connective is the first or the second arrow, and the side with two atoms has one grouping: $c_3 = c_1c_2 + c_2c_1 = 2$. For four atoms, choosing the main connective splits the chain into a left part of k atoms and a right part of $4 - k$ atoms, for $k = 1, 2, 3$, and the two parts are grouped independently:

$$c_4 = c_1c_3 + c_2c_2 + c_3c_1 = 2 + 1 + 2 = 5.$$

The three terms are the three bullets of (a).

(c) Rule 4 groups to the right, so the main connective is the first arrow and its right side is grouped to the right again: $(p \to (q \to (r \to s)))$, the first formula of the list.

The numbers $1, 1, 2, 5, 14, 42, \ldots$ produced by this splitting are the **Catalan numbers**; they grow quickly: a chain of eleven atoms joined by ten connectives already has 16,796 readings. Parentheses, or a fixed convention, are not pedantry.

4.2 Induction on the Natural Numbers

Section 4.1 made formulas into objects: strings built from atoms and $\bot$ by the clauses of Definition 4.1.1, and by nothing else. There are infinitely many of them, and the theorems this chapter is heading for are about all of them. This section supplies the tool for proving anything about infinitely many objects at once. It is **induction** on the natural numbers, where it is familiar. Section 4.3 carries it over to formulas, where the chapter needs it most.

4.2.1 One argument for infinitely many cases

Here is a claim about every formula: *every formula has as many left parentheses as right ones.* It is believable. The clause (F3) of Definition 4.1.1 adds one of each, and the other clauses add none. But the claim is about infinitely many strings, and checking them one by one never ends.

The belief already contains the shape of a proof. Every formula is reached by finitely many building steps from atoms and $\bot$ (clause (F4)). The starting pieces have no parentheses, so they have equally many of each kind. Each building step turns formulas that balance into a formula that balances. So, step by step along any formation sequence, balance is never lost, and every formula balances.

Exactly the same shape appears with numbers. To show that some property holds for every natural number, show that it holds for 0, and that it passes from each number to the next. Every natural number is reached from 0 by finitely many steps of "add one", so the property is never lost. This is **induction**. We state it for numbers first, because it is simpler there, then prove the version for formulas from it.

As in Chapters 1 and 2, "for every $n \in \mathbb{N}$" and "there is a k" are used in their everyday sense. As in the rest of the book, $\mathbb{N} = \{0, 1, 2, \ldots\}$ starts at 0, and the arithmetic and order of the natural numbers are school facts.

4.2.2 The principle of induction

A **property of natural numbers** is an open sentence $P(n)$ about a natural number n, such as "$n^2 \geq n$" or "$2^n > n$": for each particular n it becomes a statement, true or false (compare the open sentence $x > 2$ of Example (Statements and nonstatements), which became a statement once x was fixed).

If a property holds for 0, and passes from each number to the next, then it holds for every number.

Axiom 4.2.1. Principle of Induction

Let $P(n)$ be a property of natural numbers. Suppose that

(i) $P(0)$ holds, and

(ii) **for every** $n \in \mathbb{N}$, if $P(n)$ holds, then $P(n+1)$ holds.

Then $P(n)$ holds **for every** $n \in \mathbb{N}$.

In symbols (here $\forall$ abbreviates the words "for every"):

$$\Big(P(0) \ \wedge \ \forall n \in \mathbb{N}\ \big(P(n) \to P(n+1)\big)\Big) \ \to \ \forall n \in \mathbb{N}\ P(n).$$

In words: hypothesis (i) is the **base case**. Hypothesis (ii) is the **inductive step**, and inside it, the assumption "$P(n)$ holds" is the **induction hypothesis**. The step is **one** implication, proved **once**, for a number n about which we assume nothing except that it is a natural number with $P(n)$. The conclusion is about every natural number at once.

Its status. We take Axiom 4.2.1 as a basic property of $\mathbb{N}$, given and not proved. It is, in a precise sense, what the three dots in $\{0, 1, 2, \dots\}$ mean: the natural numbers are what we reach from 0 by adding one, again and again, and nothing else. Every later use cites this principle.

Dominoes. The usual picture is a row of dominoes, one for each natural number. The base case knocks over the first. The inductive step says that each domino, if it falls, knocks over the next. Then every domino falls. The picture carries the one point that matters: the step must work for **every** n, the first transition from 0 to 1 included. A single domino standing too far from the next stops the whole row, however well the others are placed.

↻ **Second reading** · *Why induction is not a rule of logic*

Not an infinite derivation. From $P(0)$ and the instances $P(0) \to P(1)$, $P(1) \to P(2)$, ... of the step we could derive $P(1)$ by $\to$E, then $P(2)$, and so on; each single $P(n)$ has a finite derivation, n lines long or so. But no finite derivation contains all of them, and "and so on" is not a rule. The principle of induction is what licenses the passage, in one step, from the two hypotheses to "for every n". It is a rule about $\mathbb{N}$, not a rule of logic.

The step as a box. In the language of Chapter 1, the inductive step is proved by a $\to$I box. We fix a natural number n, knowing nothing about it except what the step allows, and open a box assuming $P(n)$; inside the box we reach $P(n+1)$; closing the box gives the implication.

i	$P(n)$	asm, for $\to$I
	⋮	
j	$P(n+1)$	
k	$P(n) \to P(n+1)$	$\to$I i–j

For now, "let $n \in \mathbb{N}$" means what Chapter 1 made it mean (Section 1.1): one fixed number, not further specified.

Starting from another number. Often a property is claimed only from some $b \in \mathbb{N}$ on, as in "for every $n \geq 1$". Then it suffices to prove $P(b)$, and, for every $n \in \mathbb{N}$ with $n \geq b$, that $P(n)$ implies $P(n+1)$. Indeed, apply Axiom 4.2.1 to the property $Q(m)$: "$P(b+m)$ holds". Its base case $Q(0)$ is $P(b)$, and its step from $Q(m)$ to $Q(m+1)$ is the step from $P(b+m)$ to $P(b+m+1)$, where $b+m \geq b$. So $Q(m)$ holds for every $m \in \mathbb{N}$, and every natural number $n \geq b$ is $b+m$ for

the natural number $m = n - b$. We call this induction **started at** b; it is Axiom 4.2.1 applied to Q.

Here is the first proof by induction, with every part labeled.

Example. A sum of products of neighbors

Prove that for every natural number $n \geq 1$,

$$1 \cdot 2 + 2 \cdot 3 + \cdots + n(n+1) = \frac{n(n+1)(n+2)}{3}.$$

Solution. Let $P(n)$ be the displayed equation. Its left side is the sum of the n terms $k(k+1)$ for $k = 1, \ldots, n$, so the left side for $n+1$ is the left side for n plus the one new term $(n+1)(n+2)$. Before proving anything, test small cases. For $n = 1, 2, 3$ the left side is 2, $2+6 = 8$ and $8+12 = 20$, and the right side is $\frac{1\cdot2\cdot3}{3} = 2$, $\frac{2\cdot3\cdot4}{3} = 8$ and $\frac{3\cdot4\cdot5}{3} = 20$. The claim is plausible. It starts at $n = 1$, so we use induction started at 1 (see above).

Base case. For $n = 1$ the left side is the single term $1 \cdot 2 = 2$, and the right side is $\frac{1\cdot2\cdot3}{3} = 2$. So $P(1)$ holds.

Inductive step. Let $n \in \mathbb{N}$ with $n \geq 1$, and suppose $P(n)$ holds (the induction hypothesis). Then

$$\begin{aligned} 1 \cdot 2 + \cdots + n(n+1) + (n+1)(n+2) &= \frac{n(n+1)(n+2)}{3} + (n+1)(n+2) \\ &= \frac{(n+1)(n+2)(n+3)}{3}, \end{aligned}$$

where the first equality is the induction hypothesis and the second takes out the common factor $(n+1)(n+2)$, leaving $\frac{n}{3} + 1 = \frac{n+3}{3}$. The last expression is the right side of $P(n+1)$. So $P(n)$ implies $P(n+1)$.

Conclusion. By Axiom 4.2.1, started at 1, $P(n)$ holds for every natural number $n \geq 1$.

Notice where the induction hypothesis entered: exactly once, to replace the sum of the first n terms by a closed expression. A proof by induction that never uses its induction hypothesis is either not an induction or is hiding a gap.

The inductive step of an inequality usually has a second ingredient besides the hypothesis: a crude estimate that finishes the job.

Example. An inequality by induction

Prove that $3^n \geq 2n + 1$ for every $n \in \mathbb{N}$.

Solution. Let $P(n)$ be "$3^n \geq 2n + 1$".

Base case. $3^0 = 1 = 2 \cdot 0 + 1$, so $P(0)$ holds.

Inductive step. Let $n \in \mathbb{N}$ and suppose $3^n \geq 2n+1$. Then

$$3^{n+1} = 3 \cdot 3^n \geq 3(2n+1) = 6n+3 \geq 2n+3 = 2(n+1)+1,$$

where the first inequality multiplies the induction hypothesis by the positive number 3, and the second uses $6n \geq 2n$, true because $n \geq 0$. So $P(n+1)$ holds.

Conclusion. By Axiom 4.2.1, $3^n \geq 2n+1$ for every $n \in \mathbb{N}$.

The estimate $6n+3 \geq 2n+3$ throws away a great deal. That is typical: the step needs only to reach $P(n+1)$, not to reach it tightly, and a generous estimate that is obviously true is better than a sharp one that needs its own proof.

Quick check. For the claim "$5^n - 1$ is divisible by 4 for every $n \in \mathbb{N}$", write down the base case and the statement of the inductive step, without proving the step.

Solution. Base case: $5^0 - 1 = 0 = 4 \cdot 0$. Step: for every $n \in \mathbb{N}$, if $4 \mid 5^n - 1$, then $4 \mid 5^{n+1} - 1$. (It holds, since $5^{n+1} - 1 = 5(5^n - 1) + 4$.)

Warning. The inductive step must hold for **every** n, including the first transition. Try to prove "$2^n \geq n^2$ for every $n \in \mathbb{N}$": the base case $1 \geq 0$ holds, and the step $2^{n+1} = 2 \cdot 2^n \geq 2n^2 \geq (n+1)^2$ looks fine, but its last inequality holds only for $n \geq 3$, so the steps from 0 to 1, from 1 to 2 and from 2 to 3 are not covered. The claim survives the first two gaps (at $n = 1$, $2 \geq 1$, and at $n = 2$ both sides are 4), but it is false at $n = 3$, where $2^3 = 8 < 9 = 3^2$; it is true again from $n = 4$ on.

Remark. **A checklist for a proof by induction.** Before writing, and again after, answer five questions.

1. *Objects.* What does the induction run over: the natural numbers from 0, from some m, the formulas, the lines of a derivation?
2. *Property.* Write P as an open sentence about **one** object. Everything the step will need must be inside it, restrictions such as "if $n \geq 8$" included.
3. *Base cases.* Which objects does the step not reach from earlier ones: those not built from others, or, in strong induction, those whose step would reach outside the range? Check every one of them.
4. *Steps.* One for each way of building, proved for an arbitrary object, using the hypothesis only for objects built earlier: the one before, any earlier one, or the immediate parts.
5. *Principle.* Name it (Axiom 4.2.1, Theorem 4.2.2) and say what it gives.

Every induction in this chapter can be checked against these five questions, and Sections 4.3 and 4.7 do so.

4.2.3 Strong induction

Sometimes the step from n to $n+1$ is the wrong shape. The immediate parts of a formula $(\varphi \circ \psi)$ are shorter than it, but not by exactly one symbol; a natural number may be best understood through a smaller one that is not its predecessor. In such cases we want to assume the property not just for n, but for **every** smaller number.

If a property holds for n whenever it holds for all smaller numbers, then it holds for every number.

Theorem 4.2.2. Strong Induction

Let $P(n)$ be a property of natural numbers. Suppose that **for every** $n \in \mathbb{N}$: if $P(k)$ holds for **every** $k \in \mathbb{N}$ with $k < n$, then $P(n)$ holds. Then $P(n)$ holds for every $n \in \mathbb{N}$.

This form is also called **course-of-values induction**: the hypothesis may use the whole course of values $P(0), \ldots, P(n-1)$, not just the last one. Its induction hypothesis is stronger (we assume more), so its step is easier to prove; the theorem says we lose nothing by it.

There seems to be no base case. There is one, hidden. For $n = 0$, the hypothesis says: if $P(k)$ holds for every $k \in \mathbb{N}$ with $k < 0$, then $P(0)$ holds. No natural number is less than 0, so "$P(k)$ holds for every $k \in \mathbb{N}$ with $k < 0$" is **vacuously** true: for each $k \in \mathbb{N}$, the implication "if $k < 0$, then $P(k)$" has a false hypothesis, and an implication with a false hypothesis holds (Proposition 2.1.2: from $\neg(k < 0)$ we get $(k < 0) \to P(k)$). So at $n = 0$ the hypothesis of the theorem hands us $P(0)$ outright. In practice, the proof of the step usually treats small n separately, and that is where the base cases live.

Idea. Strong induction is ordinary induction applied to a stronger property. Let $Q(n)$ say "$P(k)$ holds for **every** $k \leq n$". Then $Q(n)$ is exactly the hypothesis the theorem needs at $n+1$, since the numbers below $n+1$ are those up to n; so the theorem's hypothesis moves us from $Q(n)$ to $Q(n+1)$. Ordinary induction on Q does the rest, and $Q(n)$ contains $P(n)$.

Proof. Suppose the hypothesis of the theorem holds. For $n \in \mathbb{N}$, let $Q(n)$ be the property "$P(k)$ holds for every $k \in \mathbb{N}$ with $k \leq n$". We prove $Q(n)$ for every n by Axiom 4.2.1.

Base case. No natural number is less than 0, so $P(k)$ holds for every $k \in \mathbb{N}$ with $k < 0$, vacuously. By the hypothesis of the theorem at $n = 0$, $P(0)$ holds. The only $k \in \mathbb{N}$ with $k \leq 0$ is 0, so $Q(0)$ holds.

Inductive step. Let $n \in \mathbb{N}$ and suppose $Q(n)$. The natural numbers $k < n+1$ are exactly those with $k \leq n$, since no natural number lies strictly between n and $n+1$. So $Q(n)$ says that $P(k)$ holds for every $k \in \mathbb{N}$ with $k < n+1$, and the hypothesis of the theorem at $n+1$ gives $P(n+1)$. A natural number $k \leq n+1$ satisfies $k \leq n$ or $k = n+1$, and $P(k)$ holds in both cases. Hence $Q(n+1)$.

By Axiom 4.2.1, $Q(n)$ holds for every $n \in \mathbb{N}$. Given $n \in \mathbb{N}$, $Q(n)$ applied to $k = n$ gives $P(n)$. This proves the theorem.

■

Quick check. In the proof of Theorem 4.2.2, to which property is Axiom 4.2.1 applied, and why is it not applied to $P(n)$ itself?

Solution. To $Q(n)$: "$P(k)$ holds for every $k \in \mathbb{N}$ with $k \leq n$". Applied to $P(n)$ itself, ordinary induction would give us only $P(n)$ as induction hypothesis, while the hypothesis of the theorem at $n+1$ asks for $P(k)$ at **every** $k \leq n$. The property Q carries exactly that information from one step to the next.

The first use shows the typical shape: the step splits into a few small cases, handled directly, and a general case that reaches back further than one.

Example. Threes and fives

Prove that every natural number $n \geq 8$ is a sum of threes and fives: $n = 3a + 5b$ for some integers $a, b \geq 0$.

Solution. Ordinary induction is awkward here. From $n = 3a + 5b$ we cannot easily reach $n+1$: we would have to trade some threes for fives or the other way round, and there may be none of the kind we want to trade. Going back **three** steps is easy instead: if $n-3$ is a sum of threes and fives, add one more three. So we use strong induction.

For $n \in \mathbb{N}$, let $P(n)$ be "if $n \geq 8$, then $n = 3a + 5b$ for some integers $a, b \geq 0$". Let $n \in \mathbb{N}$, and suppose $P(k)$ holds for every $k \in \mathbb{N}$ with $k < n$. We show $P(n)$.

Case $n < 8$. Then the hypothesis "$n \geq 8$" of $P(n)$ is false, and $P(n)$ holds vacuously.

Cases $n = 8, 9, 10$. Directly: $8 = 3 + 5$, $9 = 3 \cdot 3$, $10 = 5 \cdot 2$.

Case $n \geq 11$. Let $k = n - 3$. Then $k \geq 8$, so $k \in \mathbb{N}$, and $k < n$. By the induction hypothesis $P(k)$, there are integers $a, b \geq 0$ with $n - 3 = 3a + 5b$. Then $n = 3(a+1) + 5b$, with $a + 1 \geq 0$ and $b \geq 0$.

In every case $P(n)$ holds. By Theorem 4.2.2, $P(n)$ holds for every $n \in \mathbb{N}$: every $n \geq 8$ is a sum of threes and fives.

The three direct cases are the base cases. They are needed because the general case reaches back three steps: for $n = 8, 9, 10$, the number $n - 3$ is $5, 6, 7$, which is below 8, and there the induction hypothesis $P(n-3)$ is vacuous and gives us nothing. The bound 8 cannot be lowered: 7 is not a sum of threes and fives, since neither 7 nor $7 - 5 = 2$ is a multiple of 3.

4.2.4 Proving more to prove less

Sometimes the step will not go through although the claim is true. The cure is surprising: prove more.

Example. Prove more to prove less

For $n \geq 1$ let

$$S_n = \frac{1}{1 \cdot 2} + \frac{1}{2 \cdot 3} + \cdots + \frac{1}{n(n+1)}.$$

Show that $S_n < 1$ for every $n \geq 1$.

Solution. First try. Let $P(n)$ be "$S_n < 1$". The base case holds: $S_1 = \frac{1}{2}$. In the step, $S_{n+1} = S_n + \frac{1}{(n+1)(n+2)}$, and the hypothesis $S_n < 1$ tells us nothing more about S_n: a number $x < 1$ may exceed $1 - \frac{1}{(n+1)(n+2)}$, and then $x + \frac{1}{(n+1)(n+2)} > 1$. From $S_n < 1$ alone we cannot reach $S_{n+1} < 1$. The hypothesis is too weak to give $P(n+1)$. The claim is not to blame; the hypothesis is.

Guess more. $S_1 = \frac{1}{2}$, $S_2 = \frac{1}{2} + \frac{1}{6} = \frac{2}{3}$, $S_3 = \frac{2}{3} + \frac{1}{12} = \frac{3}{4}$. Let $Q(n)$ be "$S_n = \frac{n}{n+1}$".

Base case. $S_1 = \frac{1}{2} = \frac{1}{1+1}$.

Inductive step. Let $n \geq 1$ and suppose $S_n = \frac{n}{n+1}$. Then

$$S_{n+1} = \frac{n}{n+1} + \frac{1}{(n+1)(n+2)} = \frac{n(n+2)+1}{(n+1)(n+2)} = \frac{(n+1)^2}{(n+1)(n+2)} = \frac{n+1}{n+2},$$

which is $Q(n+1)$.

Conclusion. By Axiom 4.2.1 started at 1 ("Starting from another number"), $Q(n)$ holds for every $n \geq 1$, and $\frac{n}{n+1} < 1$ gives the claim.

A stronger claim gives the step a stronger hypothesis to use, and asks the step to deliver more; here it can. The same need returns with formulas in Section 4.3, where facts about parentheses are proved together, because one of them, alone, does not pass through the step.

Quick check. Run the checklist on Example (Threes and fives).

Solution. Objects: the natural numbers. Property: "if $n \geq 8$, then $n = 3a + 5b$ for some integers $a, b \geq 0$", with the restriction inside. Base cases: $8, 9, 10$, the numbers whose step would reach below 8 (for $n < 8$ the property holds vacuously). Step: for $n \geq 11$, the hypothesis for $n - 3 \geq 8$, an earlier number, and one more 3. Principle: Theorem 4.2.2.

Remark (Looking ahead). Chapter 5 gives "for every" and "there is" their rules, among them the step from "this n was arbitrary" to "for every n" that every inductive step takes.

4.2.5 Exercises

4.2.5.1 A. Check your understanding

Exercise. A1

(a) For each statement, say which kind of induction fits best (ordinary or strong), and write down its base case(s) and the statement of its inductive step, without proving them: (i) $4^n - 1$ is divisible by 3 for every $n \in \mathbb{N}$; (ii) every natural number $n \geq 12$ is $4a + 5b$ for some integers $a, b \geq 0$.

(b) True or false, with a reason: to apply Theorem 4.2.2, one must first check $P(0)$ separately.

(c) A proof of "$P(n)$ for every $n \geq 5$" checks $P(5)$ and proves "if $P(n)$, then $P(n+1)$" for every $n \geq 5$. Which form of Axiom 4.2.1 is this, and why is $P(4)$ never needed?

(d) In Example (Prove more to prove less), why can "$S_n < 1$" not serve as the induction hypothesis?

Solution. *(a)* (i) Ordinary induction. Base case: $4^0 - 1 = 0$ is divisible by 3. Step: for every $n \in \mathbb{N}$, if $3 \mid 4^n - 1$, then $3 \mid 4^{n+1} - 1$. (It holds because $4^{n+1} - 1 = 4(4^n - 1) + 3$.)

(ii) Strong induction, reaching back four: from $n - 4 = 4a + 5b$ we get $n = 4(a + 1) + 5b$. With $P(n)$ "if $n \geq 12$, then $n = 4a + 5b$ for some integers $a, b \geq 0$", the step is: for every $n \in \mathbb{N}$, if $P(k)$ holds for every $k < n$, then $P(n)$. Inside it, $n < 12$ is vacuous, and the base cases are $n = 12, 13, 14, 15$: $12 = 4 \cdot 3$, $13 = 4 \cdot 2 + 5$, $14 = 4 + 5 \cdot 2$, $15 = 5 \cdot 3$. For $n \geq 16$, $k = n - 4 \geq 12$ and $k < n$, and the hypothesis $P(n - 4)$ is used.

(b) False. At $n = 0$ the hypothesis of Theorem 4.2.2 reads "if $P(k)$ holds for every $k \in \mathbb{N}$ with $k < 0$, then $P(0)$", and its "if" part holds vacuously, since no natural number is less than 0. So proving the hypothesis for every n includes proving $P(0)$; in practice small n are treated as separate cases inside the step, as in Example (Threes and fives).

(c) Induction started at 5 ("Starting from another number"): Axiom 4.2.1 applied to $Q(m)$: "$P(m + 5)$ holds". The claim is only about $n \geq 5$, and no step reaches below 5, so nothing about 4 is ever asked.

(d) It says nothing about how close S_n is to 1: a number $x < 1$ within $\frac{1}{(n+1)(n+2)}$ of 1 has $x + \frac{1}{(n+1)(n+2)} > 1$, so from $S_n < 1$ alone the step cannot reach $S_{n+1} < 1$. The step would need a bound on how close S_n is to 1; the stronger $S_n = \frac{n}{n+1}$ gives one (and so would $S_n \leq \frac{n}{n+1}$).

Exercise. A2: Drill

The parts run from locating the hypothesis, through writing one base case or one step, to judging claims and running the checklist.

(a) In the inductive step of Example (An inequality by induction), where is the induction hypothesis used?

(b) Write down the base case and the step for "$n^2 + n$ is even for every $n \in \mathbb{N}$".

(c) Ordinary or strong induction for "every $n \geq 2$ is a product of one or more primes"?

(d) Write down $P(n + 1)$ for the property $P(n)$: "$1 + 3 + \cdots + (2n - 1) = n^2$" $(n \geq 1)$.

(e) Write down the base case for "$n! \geq 2^n$ for every $n \geq 4$" (here $n! = 1 \cdot 2 \cdots n$).

(f) Does checking $P(0)$, $P(1)$ and $P(2)$ prove the step "if $P(n)$, then $P(n + 1)$"?

(g) "$n^2 - n + 41$ is prime" holds for $n = 0, 1, \ldots, 40$. Is it true for every $n \in \mathbb{N}$?

(h) A proof shows $P(0)$ and "if $P(n)$, then $P(n+2)$" for every $n \in \mathbb{N}$. For which n is $P(n)$

proved?

(i) In Theorem 4.2.2, what may the step assume when $n = 3$?

(j) Run items 1–3 of the checklist on Example (A sum of products of neighbors).

(k) In the step for "$3 \mid n^3+2n$ for every $n \in \mathbb{N}$", $(n+1)^3+2(n+1) = (n^3+2n)+3(n^2+n+1)$. Is the hypothesis enough to finish the step?

(l) For "$S_n < 1$" in Example (Prove more to prove less), which item of the checklist failed on the first try?

Solution.

(a) In $3 \cdot 3^n \geq 3(2n+1)$.

(b) Base: $0^2 + 0 = 0$ is even. Step: if $n^2 + n$ is even, then so is $(n+1)^2 + (n+1) = (n^2+n) + 2(n+1)$.

(c) Strong: a composite $n = ab$ with $1 < a, b < n$ reaches back to a and b, not to $n-1$.

(d) $1 + 3 + \cdots + (2n-1) + (2n+1) = (n+1)^2$.

(e) $4! = 24 \geq 16 = 2^4$.

(f) No. The step is one implication for **every** n; three checked cases prove those three cases only.

(g) No: at $n = 41$ the value is $41^2 - 41 + 41 = 41^2$, not a prime. Checked cases are not a proof.

(h) For the even n only.

(i) $P(0)$, $P(1)$ and $P(2)$.

(j) Objects: the natural numbers $n \geq 1$. Property: the displayed equation for n. Base case: $n = 1$, where $1 \cdot 2 = 2 = \frac{1 \cdot 2 \cdot 3}{3}$.

(k) Yes: $n^3 + 2n$ is a multiple of 3 by hypothesis, and so is $3(n^2 + n + 1)$.

(l) Item 2: the property was too weak for the step (item 4) to go through.

4.2.5.2 B. Practice

Exercise. B1

Prove that $2^n \geq n + 1$ for every $n \in \mathbb{N}$.

Solution. Let $P(n)$ be "$2^n \geq n + 1$".

Base case. $2^0 = 1 = 0 + 1$, so $P(0)$ holds.

Inductive step. Let $n \in \mathbb{N}$ and suppose $2^n \geq n + 1$. Then

$$2^{n+1} = 2 \cdot 2^n \geq 2(n+1) = 2n + 2 \geq n + 2 = (n+1) + 1,$$

where the first inequality multiplies the induction hypothesis by $2 > 0$, and the second uses $n \geq 0$. So $P(n+1)$ holds.

By Axiom 4.2.1, $2^n \geq n + 1$ for every $n \in \mathbb{N}$.

Exercise. B2: Divisible by six

Prove by induction that $n^3 - n$ is divisible by 6 for every $n \in \mathbb{N}$. You may use that $n(n+1)$ is even for every $n \in \mathbb{N}$.

Solution. Let $P(n)$ be "$6 \mid n^3 - n$".

Base case. $0^3 - 0 = 0 = 6 \cdot 0$, so $P(0)$ holds.

Inductive step. Let $n \in \mathbb{N}$ and suppose $6 \mid n^3 - n$. Expanding,

$$(n+1)^3 - (n+1) = n^3 + 3n^2 + 2n = (n^3 - n) + 3n(n+1).$$

By the induction hypothesis, $6 \mid n^3 - n$. Since $n(n+1)$ is even, $n(n+1) = 2m$ for some integer m, and $3n(n+1) = 6m$. A sum of two multiples of 6 is a multiple of 6, so $P(n+1)$ holds.

Conclusion. By Axiom 4.2.1, $6 \mid n^3 - n$ for every $n \in \mathbb{N}$.

Marking scheme. (6 points)

- 1: states the property $P(n)$.
- 1: base case.
- 3: inductive step: the identity (1), the use of the induction hypothesis (1), the evenness of $n(n+1)$ (1).
- 1: names the principle.

Exercise. B3: Two base cases

A sequence is given by $a_0 = 2$, $a_1 = 3$, and $a_{n+2} = 3a_{n+1} - 2a_n$ for $n \in \mathbb{N}$. Prove that $a_n = 2^n + 1$ for every $n \in \mathbb{N}$. Which form of induction do you use, and why does it need two base cases?

Solution. We use Theorem 4.2.2 with $P(n)$: "$a_n = 2^n + 1$". Let $n \in \mathbb{N}$, and suppose $P(k)$ for every $k < n$.

Cases $n = 0, 1$. $a_0 = 2 = 2^0 + 1$ and $a_1 = 3 = 2^1 + 1$.

Case $n \geq 2$. Then $n-1$ and $n-2$ are natural numbers less than n, so $P(n-1)$ and $P(n-2)$ hold, and

$$a_n = 3a_{n-1} - 2a_{n-2} = 3(2^{n-1} + 1) - 2(2^{n-2} + 1) = 3 \cdot 2^{n-1} - 2^{n-1} + 1 = 2^n + 1.$$

(Check: $a_2 = 3 \cdot 3 - 2 \cdot 2 = 5 = 2^2 + 1$.)

By Theorem 4.2.2, $a_n = 2^n + 1$ for every $n \in \mathbb{N}$. The step reaches back two places, so it cannot be used at $n = 0$ or $n = 1$, where $n - 2$ is not a natural number: those two are the base cases.

Marking scheme. (6 points)

- 1: names strong induction (or ordinary induction on "$P(n)$ and $P(n+1)$").

- 2: both base cases.
- 2: the step, with the hypothesis used at $n-1$ and at $n-2$.
- 1: says why there are two base cases.

4.2.5.3 C. Going deeper

Exercise. C1: Find the error

Claim: for every $n \geq 1$,

$$1 + 2 + \cdots + n = \frac{n^2 + n + 2}{2}.$$

"Proof": suppose it holds for n. Then

$$1 + \cdots + n + (n+1) = \frac{n^2 + n + 2}{2} + (n+1) = \frac{n^2 + 3n + 4}{2} = \frac{(n+1)^2 + (n+1) + 2}{2},$$

so it holds for $n+1$; by induction it holds for every $n \geq 1$.

(a) Is the claim true for any $n \geq 1$?

(b) Which item of the checklist fails, and which part of the "proof" is correct?

(c) Show that the same step "proves" $1 + \cdots + n = \frac{n^2+n}{2} + c$ for every number c. What singles out $c = 0$?

Solution. (a) No. The true sum is $\frac{n^2+n}{2}$: that is the case $c = 0$ of (c), whose step is correct, with the base case $1 = \frac{1+1}{2}$. The claimed value is larger by 1 for every n.

(b) Item 3: no base case is checked, and at $n = 1$ it fails, since the left side is 1 and the right side $\frac{1+1+2}{2} = 2$. The step itself is correct algebra.

(c) For every c,

$$\frac{n^2 + n}{2} + c + (n+1) = \frac{n^2 + 3n + 2}{2} + c = \frac{(n+1)^2 + (n+1)}{2} + c,$$

so a correct step passes any constant along. Only the base case decides c: at $n = 1$, $1 = 1 + c$ forces $c = 0$.

Marking scheme. (10 points)

- 3: (a), with the true sum.
- 4: (b), naming the missing base case and endorsing the step.
- 3: (c).

Fatal error (0 for (b)): calling the step wrong, or claiming the formula holds at $n = 1$.

Exercise. C2: Prove more

Prove that for every $n \geq 2$,

$$\left(1 - \frac{1}{4}\right)\left(1 - \frac{1}{9}\right) \cdots \left(1 - \frac{1}{n^2}\right) > \frac{1}{2}.$$

Write R_n for this product. First show that the hypothesis $R_n > \frac{1}{2}$ alone, with $R_{n+1} = R_n(1 - \frac{1}{(n+1)^2})$, does not give $R_{n+1} > \frac{1}{2}$; then prove more.

Solution (Hint 1). Compute the product for $n = 2, 3, 4$ as fractions.

Solution (Hint 2). Write R_2, R_3, R_4 with denominators $4, 6, 8$ and guess a formula for the product; then prove that formula by induction.

Solution. The plain claim. The step multiplies by $1 - \frac{1}{(n+1)^2} < 1$, and the hypothesis $R_n > \frac{1}{2}$ tells us nothing more about R_n: for a number x just above $\frac{1}{2}$, such as $x = \frac{1}{2} + \frac{1}{4(n+1)^2}$, the number $x(1 - \frac{1}{(n+1)^2})$ is below $\frac{1}{2}$, so the hypothesis alone does not give the claim for $n + 1$.

Prove more. $R_2 = \frac{3}{4}$, $R_3 = \frac{3}{4} \cdot \frac{8}{9} = \frac{2}{3}$, $R_4 = \frac{2}{3} \cdot \frac{15}{16} = \frac{5}{8}$ suggest $R_n = \frac{n+1}{2n}$. Base case $n = 2$: $R_2 = \frac{3}{4} = \frac{2+1}{2 \cdot 2}$. Step: let $n \geq 2$ and suppose $R_n = \frac{n+1}{2n}$. Then

$$R_{n+1} = \frac{n+1}{2n} \cdot \frac{(n+1)^2 - 1}{(n+1)^2} = \frac{n+1}{2n} \cdot \frac{n(n+2)}{(n+1)^2} = \frac{n+2}{2(n+1)}.$$

By Axiom 4.2.1 started at 2, $R_n = \frac{n+1}{2n}$ for every $n \geq 2$, and so $R_n = \frac{1}{2} + \frac{1}{2n} > \frac{1}{2}$.

Marking scheme. (10 points)

- 2: argues that the plain claim does not pass through the step.
- 2: states the stronger claim.
- 2: base case.
- 3: the algebra of the step.
- 1: the conclusion, naming the principle.

Fatal error (0 for the step): "proving" the step from $R_n > \frac{1}{2}$ alone.

4.3 Induction and Recursion on Formulas

Section 4.2 proved statements about every natural number. This section proves statements about every formula in the same way, and with them answers the question Section 4.1 left open, whether every formula is read in only one way, and with it a second: does a definition with one clause per connective define exactly one thing? On the way it shows what goes wrong when an object can be built in two ways.

4.3.1 Structural induction

We return to formulas. The formation sequence of a formula plays the part that "$0, 1, 2, \ldots$" plays for numbers: every formula is reached from atoms and $\bot$ by finitely many uses of the clauses. So a property of formulas can be proved in the same two parts.

A **property of formulas** is a statement $P(\varphi)$ about a formula φ, true or false for each particular formula: "φ has as many left as right parentheses", "φ contains the atom p".

If every atom and $\bot$ has a property, and each way of building a formula passes the property from the parts to the whole, then every formula has it.

Theorem 4.3.1. Structural Induction

Let $P(\varphi)$ be a property of formulas. Suppose that

(i) $P(p)$ holds for every atom p, and $P(\bot)$ holds;

(ii) **for every** formula φ, if $P(\varphi)$ holds, then $P(\neg\varphi)$ holds;

(iii) **for all** formulas φ, ψ and **each** binary connective $\circ$, if $P(\varphi)$ and $P(\psi)$ hold, then $P((\varphi \circ \psi))$ holds.

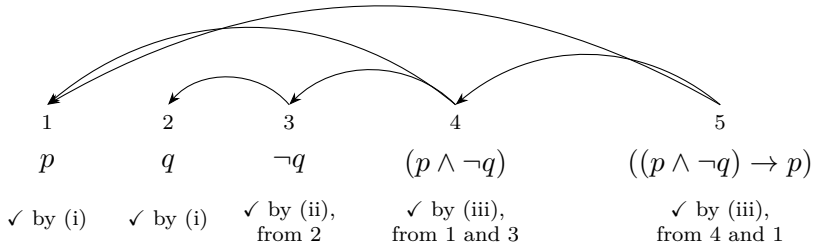
Then $P(\varphi)$ holds for every formula φ.

In words: (i) is the base case, one for each kind of formula of length one. (ii) and (iii) are the inductive steps, one for each building clause of Definition 4.1.1, and inside them "$P(\varphi)$ and $P(\psi)$" is the induction hypothesis, assumed for the **immediate parts** of the formula being built. There are seven cases in all (atoms, $\bot$, $\neg$, and the four binary connectives); often the four binary ones are handled at once, with $\circ$ left unspecified.

Section 4.2's checklist reads: objects, the formulas; property, $P(\varphi)$, with any restriction inside it; base cases, every atom and $\bot$; steps, one for $\neg$ and one for each binary connective; principle, Theorem 4.3.1.

For example, take the property "φ has as many left as right parentheses". *Base cases:* an atom and $\bot$ have no parentheses, none of either. *Step for $\neg$:* if φ has k of each, so has $\neg\varphi$, since $\neg$ is not a parenthesis. *Step for $\circ$:* if φ has k of each and ψ has l of each, then $(\varphi \circ \psi)$ has $k + l + 1$ of each, the outer pair adding one of each. By Theorem 4.3.1, every formula has as many left as right parentheses. This is part (a) of Lemma 4.3.2 below; its other parts need more care.

Idea. Take any formula and a formation sequence ending in it. Walk along the sequence: each entry is an atom or $\bot$, which has the property by (i), or is built from **earlier** entries, which have it already, by (ii) or (iii). "Earlier entries, any of them" is exactly the shape of strong induction, applied to the position in the list.



The property spreads from left to right. Each entry gets its check mark from (i), or from entries to its left that already have one.

The proof, in the fold below, makes "from left to right" exact with strong induction on the position in the list.

Second reading · *The proof of structural induction*

Proof. Suppose (i)–(iii) hold, and let χ be a formula. By Definition 4.1.1 (F4) there is a formation sequence $\chi_1, \dots, \chi_m$ with $\chi_m = \chi$. For $i \in \mathbb{N}$, let $Q(i)$ be the property "if $1 \le i \le m$, then $P(\chi_i)$ holds". We prove $Q(i)$ for every $i \in \mathbb{N}$ by Theorem 4.2.2.

Let $i \in \mathbb{N}$, and suppose $Q(k)$ holds for every $k \in \mathbb{N}$ with $k < i$. If $i = 0$ or $i > m$, then $Q(i)$ holds vacuously. So let $1 \le i \le m$. By the definition of a formation sequence, one of three things is true.

- χ_i is an atom or $\bot$. Then $P(\chi_i)$ holds by (i).
- $\chi_i = \neg\chi_j$ for some $j < i$. Then $1 \le j \le m$, and $Q(j)$ gives $P(\chi_j)$. The entry χ_j is a formula (every entry of a formation sequence is one, Section 4.1), so (ii) applies and gives $P(\chi_i)$.
- $\chi_i = (\chi_j \circ \chi_l)$ for some $j, l < i$ and a binary connective $\circ$. Then $1 \le j, l \le m$, and $Q(j)$ and $Q(l)$ give $P(\chi_j)$ and $P(\chi_l)$, both entries are formulas, and (iii) gives $P(\chi_i)$.

So $Q(i)$ holds. By Theorem 4.2.2, $Q(i)$ holds for every $i \in \mathbb{N}$; in particular $Q(m)$ gives $P(\chi_m)$, that is, $P(\chi)$. Since χ was an arbitrary formula, this proves the theorem. ■

The proof uses **strong** induction, because an entry of a formation sequence may be built from any earlier entries, not only from the one just before it.

Remark. A tempting alternative is induction on the **height** of a formula (the number of nodes on a longest branch of its parse tree, defined later in this section). That route is circular here. The height is read off the parse tree, and that a formula has only one parse tree is proved below from unique readability, whose proof uses structural induction. Inducting on the position in a formation sequence needs nothing but Definition 4.1.1.

Warning. A structural induction must cover **every** clause, the base case $\bot$ and the step for $\neg$ included. Try to prove "every formula contains an atom": it holds for atoms, and the steps for $\neg$ and for $(\varphi \circ \psi)$ keep it, since the atoms of the parts are atoms of the whole. But the case $\bot$ fails, and so does the claim: $\bot$, $\neg\bot$ and $(\bot \to \bot)$ contain no atom. A proof that "forgets" $\bot$ proves something false.

Quick check. Which clause of Definition 4.1.1 makes step (ii) of Theorem 4.3.1 necessary? Give a property of formulas that holds for every atom and for $\bot$, passes from φ and ψ to $(\varphi \circ \psi)$ for every binary connective, and yet fails for some formula.

Solution. Clause (F2), which builds $\neg\varphi$ from φ: step (ii) is the step for that clause. Without it, the formulas made by (F2) would be missed.

One such property: "φ is an atom, or is $\bot$, or begins with (". Atoms and $\bot$ have it, and every formula $(\varphi \circ \psi)$ begins with (, whatever its parts are. But $\neg p$ does not have it. So (i) and (iii) alone do not prove a property of every formula; the property fails exactly at a formula made by the forgotten clause.

Second reading · *Sets built from a base*

Section 4.1 called the formulas the **smallest** collection of strings closed under (F1)–(F3). Here is that sentence made exact.

Call a collection C of strings **closed** if it contains every atom and $\bot$, contains $\neg s$ whenever it contains s, and contains $(s \circ t)$ whenever it contains s and t, for each binary connective $\circ$. The collection of all strings is closed, and closed collections can be much larger than we want.

(i) *Every closed collection C contains every formula.* Given a formula and a formation sequence $\chi_1, \ldots, \chi_m$ ending in it, show "if $1 \leq i \leq m$, then $\chi_i \in C$" for every $i \in \mathbb{N}$ by Theorem 4.2.2, case by case as in the proof of Theorem 4.3.1: each entry is an atom or $\bot$, or is built by one clause from earlier entries, which are in C already. The last entry is the formula.

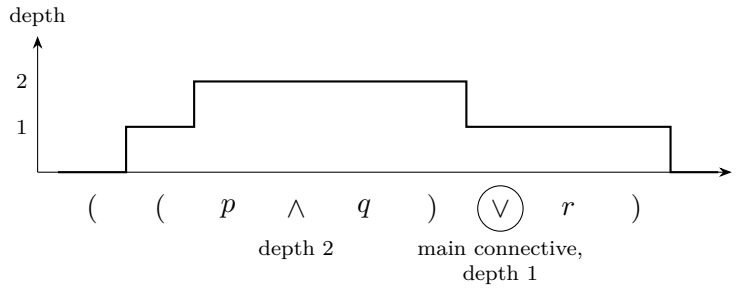
(ii) *The formulas form a closed collection*, by (F1)–(F3).

So the formulas are the smallest closed collection: contained in every closed collection, and one of them. Structural induction is (i) applied to one particular collection. If a property P satisfies (i)–(iii) of Theorem 4.3.1, the collection of formulas that have P is closed: it contains every atom and $\bot$ by (i), and the formulas built from members of it are formulas with P, by (ii) and (iii). So it contains every formula. The top-down description ("the smallest closed collection") and the bottom-up one ("whatever a formation sequence reaches") pick out the same strings. The word "collection" keeps its everyday sense here.

4.3.2 Counting parentheses: the depth profile

Section 4.1 met the string $p \wedge q \vee r$, which can be read as $(p \wedge q) \vee r$ or as $p \wedge (q \vee r)$. With p and q false and r true, the first reading is true and the second false. Section 4.4 defines the truth value of a formula one clause per connective, and the clause to use is decided by the main connective. A string with two main connectives would get two values. Formulas carry parentheses, and the parentheses are meant to settle the reading. This subsection and the next prove that they do, for every formula. The tool is a count.

Read a formula from left to right, keeping a running count: add one at each (, subtract one at each). The count just before a symbol is its **depth**: how many pairs of parentheses are open around it.



The depth profile of $((p \wedge q) \vee r)$. It starts at 0, stays at 1 or more strictly inside, comes back to 0 only after the last symbol, and the only binary connective at depth 1 is the main connective.

For every formula of the form $(\varphi \circ \psi)$ the lemma below says the same: (a) the count returns to 0, (b) it never goes below 0 and stays at 1 or more strictly inside, (c) every binary connective is at depth 1 or more. The remark after the proof adds that only the main connective is at depth exactly 1.

We count parentheses in strings, and for this some words about strings.

An **initial segment** of a string s is the string of the first k symbols of s, for some k with $0 \leq k \leq$ the length of s. For $k = 0$ it is the **empty** string, with no symbols; for k equal to the length it is s itself. An initial segment is **proper** if it is shorter than s. The **excess** of a string s is

$$\mathrm{ex}(s) := (\text{number of (in } s) - (\text{number of) in } s).$$

So a left parenthesis has excess 1, a right parenthesis -1, and any other symbol 0; and the excess of a string written as two strings one after the other is the sum of their excesses, since the counts add. For example, $\mathrm{ex}(\,(p \wedge (q\,) = 2$: after reading $(p \wedge (q$ we are inside two pairs of parentheses that have not yet closed. That is the picture to keep: the excess of an initial segment is the **depth** we are at when we have read that far. So the depth of a symbol is the excess of the initial segment that ends just before it.

Lemma 4.3.2. Parenthesis Count

Let χ be a formula.

(a) $\mathrm{ex}(\chi) = 0$: χ has as many left as right parentheses.

(b) Every initial segment u of χ has $\mathrm{ex}(u) \geq 0$. If moreover χ has the form $(\varphi \circ \psi)$ and u is proper and non-empty, then $\mathrm{ex}(u) \geq 1$.

(c) For every occurrence of a binary connective in χ, the initial segment of χ that ends just before that occurrence has excess at least 1.

In words: (a) is balance. (b) says that reading a formula from the left, we never close more parentheses than we have opened, and that inside a formula $(\varphi \circ \psi)$ we stay at depth at least 1 until the very last symbol. (c) says that no binary connective lies outside all parentheses. Part (a) is the balance proved above, repeated in the proof because (b) and (c) need it for the parts; (c), with (a), is what the unique readability theorem needs, and (b), with (a), is the key to showing that no proper initial segment of a formula is a formula (Exercise C1). The steps for (b) and (c) use (a) for the parts, so the three are proved together. Part (b)'s second sentence, alone, does not pass through the step; that is why the lemma proves more, as Example (Prove more to prove less) did (fold below).

Idea. Induction on the shape of a formula, all three parts at once. For $\neg\varphi$, the extra $\neg$ changes no count. For $(\varphi \circ \psi)$, list its initial segments by where they stop: before anything, inside φ, at the connective, inside ψ, at the end. Each is "(" followed by pieces whose excess the induction hypothesis controls.

Second reading · *The proof of the parenthesis count*

Proof. We prove, by Theorem 4.3.1, that every formula χ has the property $P(\chi)$: "(a), (b), and (c) hold for χ".

Atoms and $\bot$. The formula is one symbol, not a parenthesis. (a) Its excess is 0. (b) Its initial segments are the empty string and the formula itself, both of excess 0; and it does not have the form $(\varphi \circ \psi)$, which has length at least 5. (c) It contains no binary connective.

Negation. Let φ be a formula with $P(\varphi)$; we show $P(\neg\varphi)$. (a) $\mathrm{ex}(\neg\varphi) = 0 + \mathrm{ex}(\varphi) = 0$, by (a) for φ. (b) An initial segment of $\neg\varphi$ is either empty, of excess 0, or $\neg$ followed by an initial segment u' of φ, of excess $\mathrm{ex}(u') \geq 0$ by (b) for φ. The formula $\neg\varphi$ begins with $\neg$, so it does not have the form $(\varphi' \circ \psi')$, and the second sentence of (b) says nothing about it. (c) A binary connective in $\neg\varphi$ lies in φ, and the initial segment of $\neg\varphi$ before it is $\neg$ followed by the initial segment of φ before it, whose excess is at least 1 by (c) for φ.

Binary connectives. Let φ, ψ be formulas with $P(\varphi)$ and $P(\psi)$, let $\circ$ be a binary connective, and let $\chi = (\varphi \circ \psi)$.

(a) $\mathrm{ex}(\chi) = 1 + \mathrm{ex}(\varphi) + 0 + \mathrm{ex}(\psi) - 1 = 0$, by (a) for φ and for ψ.

(b) Every initial segment of χ is of exactly one of the following five kinds, and its excess is as shown:

- the empty string: excess 0;
- (followed by an initial segment u' of φ: excess $1 + \mathrm{ex}(u') \geq 1$, by (b) for φ;
- $(\varphi \circ$: excess $1 + \mathrm{ex}(\varphi) + 0 = 1$, by (a) for φ;
- $(\varphi \circ$ followed by a non-empty initial segment v' of ψ: excess $1 + 0 + \mathrm{ex}(v') \geq 1$, by (a) for φ and (b) for ψ;
- χ itself: excess 0, by (a) for χ.

All five have excess at least 0. The proper non-empty ones are those of the middle three kinds, and they have excess at least 1.

(c) An occurrence of a binary connective in χ is the connective $\circ$ between φ and ψ, or lies in φ, or lies in ψ. Before $\circ$ stands $(\varphi$, of excess 1. Before an occurrence in φ stands (followed by the initial segment of φ before it, of excess at least $1 + 1 = 2$ by (c) for φ. Before an occurrence in ψ stands $(\varphi \circ$ followed by the initial segment of ψ before it, of excess at least $1 + 0 + 1 = 2$ by (a) for φ and (c) for ψ. In all three cases the excess is at least 1.

So $P(\chi)$ holds. By Theorem 4.3.1, every formula has the property P, which proves the lemma.

■

The three parts were proved together for a reason. The second sentence of (b), alone, cannot be proved by structural induction. For $(\neg(p \wedge q) \wedge r)$, the initial segment $(\neg(p$ is (followed by the initial segment $\neg(p$ of the formula $\neg(p \wedge q)$, and its excess is $1 + \mathrm{ex}(\neg(p)$; to know that this is at least 1, we need $\mathrm{ex}(\neg(p) \geq 0$, a fact about an initial segment of a **negation**, which the second sentence of (b) does not cover. The first sentence of (b) does. A property that is too weak to pass through the induction must be strengthened until it does, and here the strengthening is (b)'s first sentence, as the strengthening of Example (Prove more to prove less) was $S_n = \frac{n}{n+1}$.

The proof (in the fold) gives a little more than (c) states, and the next proof uses the extra. In a formula $(\varphi \circ \psi)$, the initial segment before the connective $\circ$ has excess exactly 1, and the initial segment before **any other** binary connective has excess at least 2. In the language of depth: the main connective is the only binary connective at depth one. That is the recipe of Section 4.1 for finding the root of a parse tree.

Quick check. Write down the depth of each symbol of $(p \wedge (q \vee r))$. Which binary connective is at depth 1?

Solution. (0, p 1, $\wedge$ 1, (1, q 2, $\vee$ 2, r 2,) 2,) 1, and 0 after the last symbol. The $\wedge$ is at depth 1, and it is the main connective. Compare the figure, whose main connective was the $\vee$.

4.3.3 Unique readability

We can now prove what Section 4.1 promised: a formula can be built in only one way.

Theorem 4.3.3. Unique Readability

Let χ be a formula that is neither an atom nor $\bot$. Then **exactly one** of the following holds:

(i) $\chi = \neg\varphi$ for a formula φ;
(ii) $\chi = (\varphi \circ \psi)$ for formulas φ, ψ and a binary connective $\circ$.

In case (i), φ is **uniquely** determined by χ; in case (ii), $\circ$, φ and ψ are uniquely determined by χ.

In words: every compound formula has one main connective, and one way of being split into its immediate parts. This is the precise form of Chapter 1's claim that a statement built from letters by connectives can be read in only one way. Existence is the easy half; the content is **uniqueness**. The danger it excludes is a string such as $p \wedge q \vee r$, which, if it were a formula, could be split at $\wedge$ or at $\vee$.

Idea. Some form holds, by looking at the last step (Section 4.1); the first symbol tells the two forms apart. For case (ii), suppose $\chi = (\varphi \circ \psi) = (\varphi' \circ' \psi')$. By the remark after Lemma 4.3.2, each main connective is the only binary connective at depth one; so both sit at the same position of the string, and everything else is forced.

Proof. By Section 4.1 (looking at the last step), χ is an atom, $\bot$, $\neg\varphi$ for a formula φ, or $(\varphi \circ \psi)$ for formulas φ, ψ; by hypothesis it is one of the last two. It cannot be both, since a string of the form $\neg\varphi$ begins with $\neg$ and one of the form $(\varphi \circ \psi)$ begins with (.

Case (i). If $\chi = \neg\varphi = \neg\varphi'$, removing the first symbol gives $\varphi = \varphi'$.

Case (ii). Let $\chi = (\varphi \circ \psi)$. Before the connective $\circ$ stands the initial segment $(\varphi$, of excess $1 + \mathrm{ex}(\varphi) = 1$ by Lemma 4.3.2 (a). Before a binary connective inside φ stands (followed by an initial segment of φ of excess at least 1, by Lemma 4.3.2 (c); so the excess is at least 2. Before a binary connective inside ψ stands $(\varphi \circ$ followed by an initial segment of ψ of excess at least 1, by (c) for ψ and (a) for φ; so again at least 2. Hence the position of $\circ$ in χ is **the only** position of a binary connective in χ whose initial segment before it has excess 1, and this description mentions only the string χ.

Now suppose also $\chi = (\varphi' \circ' \psi')$ for formulas φ', ψ' and a binary connective $\circ'$. By the same argument, the position of $\circ'$ is the only position with that description. So $\circ$ and $\circ'$ stand at the same position of χ, and $\circ = \circ'$. That position is the one just after (and φ, and also the one just after (and φ', so φ and φ' have the same length; both are the string of symbols of χ between the first symbol and that position, so $\varphi = \varphi'$. Likewise ψ and ψ' are both the string of symbols between that position and the last symbol, so $\psi = \psi'$. This proves the theorem. ■

The theorem settles the question Section 4.1 left open about parse trees.

Every formula has exactly one parse tree. We know it has one. That it has only one is proved by structural induction. By (T4) of Definition 4.1.2, every parse tree is made by one of (T1)–(T3). An atom or $\bot$ has only the one-node tree, since a tree made by (T2) or (T3) of Definition 4.1.2 is a tree for a string beginning with $\neg$ or (. A parse tree for $\neg\varphi$ must have been made by (T2), from a parse tree for some φ' with $\neg\varphi' = \neg\varphi$; so $\varphi' = \varphi$, and by the induction hypothesis the tree below the root is the one parse tree of φ. A parse tree for $(\varphi \circ \psi)$ must have been made by (T3), from parse trees for some φ', ψ' and a connective $\circ'$ with $(\varphi' \circ' \psi') = (\varphi \circ \psi)$; by Theorem 4.3.3, $\circ' = \circ$, $\varphi' = \varphi$ and $\psi' = \psi$, and by the induction hypothesis the two subtrees are the parse trees of φ and ψ. From now on we speak of **the** parse tree of a formula.

With one parse tree for each formula, a measurement of that tree is a measurement of the formula.

Definition 4.3.1. Height

The **height** of a formula φ, written $\mathrm{ht}(\varphi)$, is the number of nodes on a **longest** branch of its parse tree.

Atoms and $\bot$ have height 1: their tree is one node, and its only branch has one node. (Some books count edges instead of nodes and give atoms height 0; with nodes, every formula has height at least 1.) In Example (Three parse trees), (b) has height 4, along the branch $\neg, \wedge, \neg, q$, and (c) has height 4 too, along $\vee, \leftrightarrow, \neg, q$ or along $\vee, \wedge, \rightarrow, \bot$.

Warning. The height counts the nodes on the longest **branch**, not the symbols of the formula. The formula $\neg(p \wedge \neg q)$ has seven symbols, $\neg$, (, p, $\wedge$, $\neg$, q,), and height 4; the formula $((p \wedge q) \wedge (r \wedge p))$ has thirteen symbols and height only 3. Parentheses never count

towards the height, and a wide tree can be shallow.

4.3.4 Definition by recursion

Height, the set of atoms, the number of connectives: each is naturally defined **one clause per way of building a formula**. The height of $\neg\varphi$ is one more than that of φ; the atoms of $(\varphi \circ \psi)$ are those of φ together with those of ψ. Such a definition is called **recursive**: it defines the value at a formula through the values at its immediate parts. The question is whether it really defines something. For a string that could be split in two ways, the clauses might assign two different values, and then they would define nothing. Unique readability says this never happens.

We say **function on formulas** for a rule that assigns to each formula exactly one value; the values may be numbers, truth values, formulas, or collections of atoms. Nothing here needs more than this everyday sense of "function".

Give a value for each atom and for $\bot$, and say how to get the value of a compound formula from its parts; then exactly one function on formulas obeys these clauses.

Theorem 4.3.4. Structural Recursion

Suppose we are given

- a value a_p for each atom p, and a value $a_\bot$;
- a rule $g_\neg$ that produces a value $g_\neg(\varphi, x)$ from any formula φ and any value x;
- for each binary connective $\circ$, a rule $g_\circ$ that produces a value $g_\circ(\varphi, \psi, x, y)$ from any formulas φ, ψ and any values x, y.

Then there is **exactly one** function f on formulas such that, for every atom p, all formulas φ, ψ and every binary connective $\circ$,

$$\begin{aligned} f(p) &= a_p, \qquad f(\bot) = a_\bot, \\ f(\neg\varphi) &= g_\neg(\varphi, f(\varphi)), \\ f((\varphi \circ \psi)) &= g_\circ(\varphi, \psi, f(\varphi), f(\psi)). \end{aligned}$$

In words: the three lines are the **clauses** of a recursive definition, one for each clause of Definition 4.1.1. The rules $g_\neg$ and $g_\circ$ may look at the immediate parts themselves as well as at their values; most definitions use only the values, and the list of subformulas below uses the parts. The theorem says the clauses define a function (**existence**) and only one (**uniqueness**).

Idea. Build the function on short formulas first. A formula with one symbol is an atom or $\bot$, and the first clauses fix its value. A longer formula has immediate parts that are shorter, whose values are already fixed, and unique readability says which clause applies and to which parts, with no choice left. So the value at each formula is forced, length by length. Formally: an ordinary induction on n, proving that there is exactly one function on the formulas of length at most n that obeys the clauses.

Proof. For $n \in \mathbb{N}$ with $n \geq 1$, call a function h, defined on the formulas of length at most n, **n-good** if it obeys the three clauses at every formula of length at most n. This makes sense: the immediate parts of a formula are shorter than it, so the clauses at a formula of length at most n mention values of h only at formulas of length at most n. Let $S(n)$ be "there is exactly one n-good function"; we prove $S(n)$ for every $n \geq 1$ by Axiom 4.2.1, started at 1 ("Starting from another number", Section 4.2).

Base case. By the first-symbol test (Section 4.1), the formulas of length 1 are the atoms and $\bot$. A 1-good function must give each atom p the value a_p and $\bot$ the value $a_\bot$, and the function that does so is 1-good, since the other two clauses concern no formula of length 1. So $S(1)$ holds.

Inductive step. Let $n \in \mathbb{N}$ with $n \geq 1$, suppose $S(n)$, and let h_n be the n-good function. Define h on the formulas of length at most $n+1$ as follows. At a formula of length at most n, let h agree with h_n. Let χ have length $n+1 \geq 2$; it is neither an atom nor $\bot$, so by Theorem 4.3.3 it is **either** $\neg\varphi$ for exactly one formula φ, **or** $(\varphi \circ \psi)$ for exactly one $\circ$, φ, ψ; the parts have length at most n. Let $h(\chi) \coloneqq g_\neg(\varphi, h_n(\varphi))$ in the first case and $h(\chi) \coloneqq g_\circ(\varphi, \psi, h_n(\varphi), h_n(\psi))$ in the second. Unique readability leaves no choice of case or of parts, so this is one value. The function h is $(n+1)$-good: at formulas of length at most n it agrees with the n-good h_n, including at their parts, and at formulas of length $n+1$ it obeys the clauses by construction.

For uniqueness, let h' be any $(n+1)$-good function. Its restriction to the formulas of length at most n is n-good, so it is h_n, by $S(n)$. At a formula $\chi = \neg\varphi$ of length $n+1$, the clause for $\neg$ gives $h'(\chi) = g_\neg(\varphi, h'(\varphi)) = g_\neg(\varphi, h_n(\varphi)) = h(\chi)$; at $\chi = (\varphi \circ \psi)$ the same argument with $g_\circ$ gives $h'(\chi) = h(\chi)$. So $h' = h$, and $S(n+1)$ holds. By Axiom 4.2.1, started at 1, $S(n)$ holds for every $n \geq 1$; let h_n denote the n-good function.

The function f. If $m \geq n$, the restriction of h_m to the formulas of length at most n is n-good, so it is h_n. For a formula χ of length n, let $f(\chi) \coloneqq h_n(\chi)$; then $f(\chi) = h_m(\chi)$ for every $m \geq n$. The function f obeys the clauses: for $\chi = \neg\varphi$ of length m, $f(\chi) = h_m(\chi) = g_\neg(\varphi, h_m(\varphi)) = g_\neg(\varphi, f(\varphi))$, and likewise for the other two clauses. And it is the only one: a function f' on formulas obeying the clauses restricts, for each n, to an n-good function, which is h_n; so $f'(\chi) = h_n(\chi) = f(\chi)$ for a formula χ of length n. This proves the theorem. ■

Unique readability was used at one place: to define $h(\chi)$ by exactly one clause, applied to exactly one choice of parts. Without it, the clauses could demand two different values at one string, and then no function would obey them.

What goes wrong without it. The positive integers are built from 1 by two steps, "add one" and "double"; adding one alone reaches every positive integer. Try to define a function one clause per building step: $g(1) \coloneqq 0$, $g(n+1) \coloneqq g(n) + 1$, $g(2n) \coloneqq g(n)$. The number 2 is built in two ways, as $1+1$ and as twice 1. The first clause gives $g(2) = g(1) + 1 = 1$, the second $g(2) = g(1) = 0$, and no function obeys both. The positive integers are not built **freely** by these two steps: one number can come out of two different steps, which is exactly what Theorem 4.3.3 rules out for formulas.

Quick check. Do the clauses $h(1) := 1$, $h(n+1) := h(n)+2$, $h(2n) := 2h(n)$ define a function on the positive integers?

Solution. No. At 2, the first step clause gives $h(2) = h(1) + 2 = 3$ and the second $h(2) = 2h(1) = 2$.

Example. Four recursive definitions

Each of the following is a function on formulas by Theorem 4.3.4.

- **Height:** $\mathrm{ht}(p) = \mathrm{ht}(\bot) = 1$, $\mathrm{ht}(\neg\varphi) = \mathrm{ht}(\varphi) + 1$, and
$$\mathrm{ht}((\varphi \circ \psi)) = 1 + \max\{\mathrm{ht}(\varphi), \mathrm{ht}(\psi)\}.$$
- **Atoms:** $\mathrm{atoms}(p) = \{p\}$; $\mathrm{atoms}(\bot)$ has no members; $\mathrm{atoms}(\neg\varphi) = \mathrm{atoms}(\varphi)$; and $\mathrm{atoms}((\varphi \circ \psi))$ consists of the members of $\mathrm{atoms}(\varphi)$ together with those of $\mathrm{atoms}(\psi)$.
- **Number of connectives:** 0 for atoms and $\bot$; one more than for φ at $\neg\varphi$; one more than the sum for φ and ψ at $(\varphi \circ \psi)$.
- **Subformulas:** the list of subformulas of p is p, and of $\bot$ is $\bot$; of $\neg\varphi$, the formula $\neg\varphi$ followed by the subformulas of φ; of $(\varphi \circ \psi)$, the formula $(\varphi \circ \psi)$ followed by the subformulas of φ and then those of ψ, each formula kept only at its first appearance.

Compute all four for $\chi = (\neg(p \wedge q) \to (q \vee \bot))$, and also its length.

Solution. Work up from the leaves. For the parts: $\mathrm{ht}(p \wedge q) = 1 + \max\{1, 1\} = 2$, so $\mathrm{ht}(\neg(p \wedge q)) = 3$; $\mathrm{ht}(q \vee \bot) = 2$. Hence
$$\mathrm{ht}(\chi) = 1 + \max\{3, 2\} = 4.$$
The atoms of $\neg(p \wedge q)$ are p and q, those of $q \vee \bot$ are just q ($\bot$ contributes none), so $\mathrm{atoms}(\chi) = \{p, q\}$. The connectives: $\wedge$ and $\neg$ on the left, $\vee$ on the right, $\to$ at the root, four in all. The subformulas, in the order the clause lists them: χ, $\neg(p \wedge q)$, $(p \wedge q)$, p, q, $(q \vee \bot)$, $\bot$; the second q is dropped, being already listed. The length is 14: the official formula consists of the fourteen symbols
$$(\ \neg\ (\ p\ \wedge\ q\ )\ \to\ (\ q\ \vee\ \bot\ )\ ).$$

Two checks connect these definitions with the parse trees. The **recursive height agrees with the tree height**: the parse tree of $\neg\varphi$ is that of φ with one node on top, so each branch gains one node; the tree of $(\varphi \circ \psi)$ is a new root over the trees of φ and ψ, so a longest branch is the root followed by a longest branch of the taller of the two. A structural induction whose cases are these two sentences (and the one-node tree for atoms and $\bot$) shows the two heights are equal. In the same way, the recursive list of subformulas is the list of formulas at the nodes of the parse tree (Definition 4.1.3): the nodes of the tree of $\neg\varphi$ are its root, carrying $\neg\varphi$, and the nodes of the tree of φ.

The length $\mathrm{len}(\varphi)$ of a formula φ, its number of symbols, also satisfies clauses of this shape, by counting symbols: a formula of length one for atoms and $\bot$, one more at $\neg\varphi$, and at $(\varphi \circ \psi)$ the lengths of φ and ψ plus 3 (two parentheses and the connective). For χ above: $(p \wedge q)$ has

$1 + 1 + 3 = 5$ symbols, so $\neg(p \wedge q)$ has 6; $(q \vee \bot)$ has 5; so χ has $6 + 5 + 3 = 14$, as counted.

Warning. A recursive definition must have a clause for **every** way of building a formula, and each clause may use only the **immediate parts**. "Let $f(\varphi)$ be f of the longest subformula of φ plus one" is not a recursive definition: at φ itself the longest subformula is φ, and the "definition" refers to the value it is trying to define. The theorem applies only to clauses of the form it states.

Remark (Looking ahead). Chapter 5 builds the formulas of predicate logic by clauses of the same kind, and proves things about all of them by inductions of this kind.

4.3.5 Exercises

4.3.5.1 A. Check your understanding

Exercise. A1

(a) Which kind of induction fits "every formula contains an atom or $\bot$"? Write down its base cases and the statements of its steps, without proving them.

(b) A reader proves "every formula has at least one symbol that is not a parenthesis" by structural induction. They check atoms, the step for $\neg$, and the step for $(\varphi \circ \psi)$. Which case is missing? Does the claim hold there?

(c) Write down, without proving them, the base cases and the inductive steps of a structural induction for: every formula has exactly as many left parentheses as binary connectives.

(d) Which parts of Lemma 4.3.2 does the proof of Theorem 4.3.3 use, and for which formulas?

(e) Why does Theorem 4.3.4 not apply to the positive integers built from 1 by "add one" and "double"?

Solution. *(a)* Structural induction. Base cases: an atom contains itself, and $\bot$ contains $\bot$. Steps: if φ contains an atom or $\bot$, so does $\neg\varphi$; if φ and ψ do, so does $(\varphi \circ \psi)$, for each binary $\circ$. (Compare the warning in the section: with "an atom" alone, the case $\bot$ fails.)

(b) The case $\bot$ is missing. The claim holds there, since $\bot$ is itself such a symbol. The proof is incomplete, though the claim is true; compare the warning on "every formula contains an atom", where the missing case makes the claim false.

(c) Base cases: an atom and $\bot$ have none of either. Step for $\neg$: if φ has as many of each, so has $\neg\varphi$, which adds neither. Step for $(\varphi \circ \psi)$: if each of φ, ψ has as many of each, the whole adds one left parenthesis and one binary connective.

(d) Part (a) for φ, to get excess 1 before the main connective and to get excess $1+0$ before the part ψ; part (c) for φ and for ψ, to get excess at least 2 before every other binary connective. Here $\chi = (\varphi \circ \psi)$ is the formula being read. Part (b) is not used in that proof.

(e) Its proof needs each compound object to be built in exactly one way, which for formulas is unique readability. The number 2 is built as $1 + 1$ and as twice 1, so clauses can demand

two values at one number. Some clauses happen to agree wherever a number is built twice: $g(1) := 1$, $g(n+1) := g(n)+1$, $g(2n) := 2g(n)$ is obeyed by $g(n) = n$. But without unique building nothing guarantees this in advance.

Exercise. A2: Drill

The parts run from naming the induction, through computing depths, heights and recursive values, to judging claims and definitions and running the checklist.

(a) Which kind of induction proves "every formula has as many left as right parentheses"?

(b) Write down the depth of each symbol of $\neg(p \to \neg q)$. Which binary connective is at depth 1?

(c) What is the excess of the initial segment $((p \wedge (q$ of $((p \wedge (q \vee r)) \to p)$?

(d) Find the main connective of $((p \to q) \wedge (r \vee p))$ by depth.

(e) What is the height of $((p \wedge q) \vee \neg\neg r)$?

(f) True or false: every subformula of φ other than φ itself has smaller height than φ.

(g) Compute the recursive height and the recursive list of subformulas of $\neg(p \vee \neg p)$.

(h) Define by recursion the number $b(\varphi)$ of occurrences of $\bot$ in φ, and compute $b((\bot \to \neg\bot))$.

(i) Is "$f(\varphi) := f(\varphi) + 1$ for every formula φ" a definition by recursion?

(j) "Every formula has odd length or begins with $\neg$." Which case of a structural induction fails, and on which formula does the claim fail?

(k) Do the clauses $g(1) := 1$, $g(n+1) := g(n)+1$, $g(2n) := g(n)+1$ define a function on the positive integers built by "add one" and "double"?

(l) Run the checklist on "every formula with no $\neg$ and no $\bot$ has odd length".

Solution.

(a) Structural induction, or strong induction on the position in a formation sequence, which is how structural induction is proved.

(b) $\neg$ 0, (0, p 1, $\to$ 1, $\neg$ 1, q 1,) 1, and 0 after. The $\to$ is at depth 1: it is the main connective of the part inside the parentheses. The main connective of the whole formula is the first $\neg$.

(c) 3: three left parentheses, no right one.

(d) The $\wedge$, the only binary connective at depth 1; the $\to$ and the $\vee$ are at depth 2.

(e) $1 + \max\{2, 3\} = 4$.

(f) True. A subformula other than φ sits at a node below the root, and the part of the tree below that node is its parse tree. A longest branch of that part, with the path from the root of φ's tree down to the node added in front, is a branch of φ's tree with at least one more node. So φ's tree has a branch longer than every branch of the subformula's tree.

(g) Height 4 (p: 1, $\neg p$: 2, $(p \vee \neg p)$: 3). List: $\neg(p \vee \neg p)$, $(p \vee \neg p)$, p, $\neg p$.

(h) $b(p) = 0$ for every atom p, $b(\bot) = 1$, $b(\neg\varphi) = b(\varphi)$, $b((\varphi \circ \psi)) = b(\varphi) + b(\psi)$. Value $1 + 1 = 2$.

(i) No. It gives no value at atoms or $\bot$, and it defines the value at φ from the value at

φ itself, not from the immediate parts. No function obeys it, since no number equals itself plus one.

(j) The step for $(\varphi \circ \psi)$, whose length is $\text{len}(\varphi) + \text{len}(\psi) + 3$: the hypothesis allows a part that begins with $\neg$ and has even length. The claim fails on $(p \wedge \neg q)$, of length 6, beginning with (.

(k) No: $g(2) = 2$ and $g(3) = 3$ by the first step clause, and then $g(4) = g(3) + 1 = 4$ by the first step clause but $g(4) = g(2) + 1 = 3$ by the second.

(l) Objects: formulas. Property: "if φ contains no $\neg$ and no $\bot$, then its length is odd", the restriction inside. Base cases: an atom (length 1), and $\bot$ (vacuous). Steps: $\neg\varphi$ is vacuous; for $(\varphi \circ \psi)$ without $\neg$ and $\bot$, its parts satisfy the restriction, so their lengths are odd, and odd + odd +3 is odd. Principle: Theorem 4.3.1. (Compare (j): with $\neg$ allowed the claim fails.)

4.3.5.2 B. Practice

Exercise. B1

Prove by structural induction that every formula has at most as many distinct subformulas as it has symbols. Use the recursive description of the subformulas in Example (Four recursive definitions) and the clauses for the length given after it.

Solution. For a formula φ, write $s(\varphi)$ for the number of distinct subformulas of φ. We prove $s(\varphi) \le \text{len}(\varphi)$ by Theorem 4.3.1.

Atoms and $\bot$. The only subformula is the formula itself: $s = 1 = \text{len}$.

Negation. Suppose $s(\varphi) \le \text{len}(\varphi)$. The subformulas of $\neg\varphi$ are $\neg\varphi$ and those of φ, so there are at most $1 + s(\varphi)$ of them (fewer if one repeats). Hence $s(\neg\varphi) \le 1 + s(\varphi) \le 1 + \text{len}(\varphi) = \text{len}(\neg\varphi)$.

Binary connectives. Suppose $s(\varphi) \le \text{len}(\varphi)$ and $s(\psi) \le \text{len}(\psi)$. The subformulas of $(\varphi \circ \psi)$ are the formula itself and those of φ and of ψ, so

$$s((\varphi \circ \psi)) \le 1 + s(\varphi) + s(\psi) \le 1 + \text{len}(\varphi) + \text{len}(\psi) < \text{len}(\varphi) + \text{len}(\psi) + 3,$$

and the right side is $\text{len}((\varphi \circ \psi))$.

By Theorem 4.3.1, $s(\varphi) \le \text{len}(\varphi)$ for every formula φ. (The bound is attained by p, $\neg p$, $\neg\neg p$, and every formula built from an atom or $\bot$ by negations alone; each binary connective adds three symbols but at most one new subformula, so the inequality is strict for every formula that contains a binary connective.)

Exercise. B2

The number of nodes of the parse tree of a formula is given by the recursion $\text{nodes}(p) = \text{nodes}(\bot) = 1$, $\text{nodes}(\neg\varphi) = \text{nodes}(\varphi) + 1$, $\text{nodes}((\varphi \circ \psi)) = \text{nodes}(\varphi) + \text{nodes}(\psi) + 1$. Prove that every formula φ of height h satisfies $\text{nodes}(\varphi) \le 2^h - 1$, and give a formula of height 3 for which equality holds.

Solution. We prove $\text{nodes}(\varphi) \le 2^{\text{ht}(\varphi)} - 1$ by Theorem 4.3.1, using the recursive height of Example (Four recursive definitions).

Atoms and $\bot$. $\text{nodes} = 1 = 2^1 - 1$.

Negation. Let $h = \text{ht}(\varphi)$ and suppose $\text{nodes}(\varphi) \le 2^h - 1$. Then $\text{ht}(\neg\varphi) = h + 1$ and

$$\text{nodes}(\neg\varphi) \le 2^h - 1 + 1 = 2^h \le 2^{h+1} - 1,$$

where the last step is $1 \le 2^h$.

Binary connectives. Let $a = \text{ht}(\varphi)$, $b = \text{ht}(\psi)$, and suppose the bound for both. Then $h \coloneqq \text{ht}((\varphi \circ \psi)) = 1 + \max\{a, b\}$, so $a \le h - 1$ and $b \le h - 1$, and $2^a, 2^b \le 2^{h-1}$. Hence

$$\text{nodes}((\varphi \circ \psi)) \le (2^a - 1) + (2^b - 1) + 1 \le 2 \cdot 2^{h-1} - 1 = 2^h - 1.$$

By Theorem 4.3.1, the bound holds for every formula. Equality at height 3: $((p \wedge q) \wedge (r \wedge p))$ has $7 = 2^3 - 1$ nodes (a root, two $\wedge$ nodes, and four leaves). Equality needs every node above the bottom level to have two children.

Exercise. B3: The mirror image of a formula

Define a function m on formulas by recursion: $m(p) = p$ for every atom p, $m(\bot) = \bot$, $m(\neg\varphi) = \neg m(\varphi)$, and $m((\varphi \circ \psi)) = (m(\psi) \circ m(\varphi))$ for each binary connective $\circ$.

(a) Compute $m((p \to (q \wedge \neg r)))$.

(b) Prove that $m(m(\varphi)) = \varphi$ for every formula φ.

(c) Prove that $\text{ht}(m(\varphi)) = \text{ht}(\varphi)$ for every formula φ.

(d) Give a formula φ, not an atom or $\bot$, with $m(\varphi) = \varphi$.

Solution. *(a)* $m((p \to (q \wedge \neg r))) = (m((q \wedge \neg r)) \to p) = ((\neg r \wedge q) \to p)$.

(b) By Theorem 4.3.1, with $P(\varphi)$: "$m(m(\varphi)) = \varphi$". *Atoms and $\bot$:* m fixes them, so $m(m(p)) = p$ and $m(m(\bot)) = \bot$. *Negation:* if $m(m(\varphi)) = \varphi$, then $m(m(\neg\varphi)) = m(\neg m(\varphi)) = \neg m(m(\varphi)) = \neg\varphi$. *Binary connectives:* if $m(m(\varphi)) = \varphi$ and $m(m(\psi)) = \psi$, then

$$m(m((\varphi \circ \psi))) = m((m(\psi) \circ m(\varphi))) = (m(m(\varphi)) \circ m(m(\psi))) = (\varphi \circ \psi).$$

(c) By Theorem 4.3.1, with the recursive height of Example (Four recursive definitions). *Atoms and $\bot$:* both sides are 1. *Negation:* $\text{ht}(m(\neg\varphi)) = \text{ht}(\neg m(\varphi)) = \text{ht}(m(\varphi)) + 1 = \text{ht}(\varphi) + 1 = \text{ht}(\neg\varphi)$. *Binary connectives:*

$$\text{ht}(m((\varphi \circ \psi))) = 1 + \max\{\text{ht}(m(\psi)), \text{ht}(m(\varphi))\} = 1 + \max\{\text{ht}(\varphi), \text{ht}(\psi)\} = \text{ht}((\varphi \circ \psi)),$$

using the hypothesis for both parts.

(d) $(p \wedge p)$, or $\neg(q \vee q)$.

Marking scheme. (8 points)

- 1: (a).

- 3: (b), all four cases (atoms, $\bot$, negation, binary).
- 3: (c), a structural induction (on the recursive height, or showing that the tree of $m(\varphi)$ mirrors that of φ) using the hypothesis for both parts.
- 1: (d).

4.3.5.3 C. Going deeper

Exercise. C1: Initial segments are not formulas

(a) Prove that no proper initial segment of a formula is a formula.

(b) Deduce, in a way different from the proof in the text, the uniqueness part of case (ii) of Theorem 4.3.3: if $(\varphi \circ \psi) = (\varphi' \circ' \psi')$ for formulas $\varphi, \psi, \varphi', \psi'$ and binary connectives $\circ, \circ'$, then $\varphi = \varphi'$, $\circ = \circ'$ and $\psi = \psi'$.

Solution (Hint 1). Structural induction on the formula whose initial segments you consider. Which part of Lemma 4.3.2 settles the case $(\varphi \circ \psi)$ at once?

Solution (Hint 2). For (b): remove the first symbol from both sides. Now both φ and φ' are initial segments of the same string.

Solution. (a) The empty string is not a formula, since every formula has at least one symbol. So it is enough to show that no proper **non-empty** initial segment of a formula χ is a formula. We prove this by Theorem 4.3.1.

Atoms and $\bot$. A formula of length 1 has no proper non-empty initial segment.

Negation. Suppose no proper non-empty initial segment of φ is a formula, and let u be a proper non-empty initial segment of $\neg\varphi$. Then u is $\neg$ followed by a proper initial segment u' of φ. Suppose, for a contradiction, that u is a formula. It begins with $\neg$, so it is not an atom or $\bot$ and not of the form $(\varphi'' \circ \psi'')$; by looking at the last step (Section 4.1), $u = \neg\varphi''$ for a formula φ'', and removing the first symbol gives $u' = \varphi''$. So u' is a formula, hence non-empty, and it is a proper non-empty initial segment of φ, contradicting the induction hypothesis.

Binary connectives. Let u be a proper non-empty initial segment of $(\varphi \circ \psi)$. By Lemma 4.3.2 (b), $\mathrm{ex}(u) \geq 1$. If u were a formula, Lemma 4.3.2 (a) would give $\mathrm{ex}(u) = 0$. So u is not a formula. (This case needs no induction hypothesis.)

By Theorem 4.3.1, no formula has a proper non-empty initial segment that is a formula, which proves (a).

(b) Suppose $(\varphi \circ \psi) = (\varphi' \circ' \psi')$. Removing the first symbol, the strings $\varphi \circ \psi)$ and $\varphi' \circ' \psi')$ are equal. Both φ and φ' are initial segments of this string, so the shorter of the two is an initial segment of the longer. If their lengths differed, the shorter would be a proper initial segment of the longer, and a formula, contradicting (a). So they have the same length, and being initial segments of the same string, $\varphi = \varphi'$. The next symbol is then $\circ$ on one side and $\circ'$ on the other, so $\circ = \circ'$; what remains is $\psi) = \psi')$, and removing the last symbol gives $\psi = \psi'$.

Marking scheme. (10 points)

- 1: reduces (a) to non-empty segments (the empty string is not a formula).
- 2: sets up a structural induction (or a strong induction on the length) with the property "no proper non-empty initial segment is a formula", with the atom and $\bot$ case.
- 3: negation case: reads the segment as $\neg$ followed by a proper initial segment of φ, and uses the last step of a formula beginning with $\neg$ to reach a contradiction with the hypothesis.
- 2: binary case: excess at least 1 by Lemma 4.3.2 (b) against excess 0 by (a).
- 2: (b): compares φ and φ' as initial segments of one string, concludes equal length by (a), then reads off $\circ$ and ψ.

Fatal errors (0 for the step): using Theorem 4.3.3 to prove (b); assuming the lengths of φ and φ' are equal.

Exercise. C2: Find the error

The following argument claims to prove that every formula contains an **even** number of parentheses.

(1) We prove, by strong induction on n, that every formula with n symbols has an even number of parentheses.

(2) Let $n \in \mathbb{N}$, let χ be a formula with n symbols, and suppose that every formula with fewer than n symbols has an even number of parentheses.

(3) If χ is an atom or $\bot$, it has no parentheses, and 0 is even.

(4) If $\chi = \neg\varphi$, then φ is a formula with $n-1$ symbols, so by hypothesis it has an even number of parentheses; and χ has the same parentheses as φ.

(5) If $\chi = (\varphi \circ \psi)$, remove the outer pair of parentheses. The string $\varphi \circ \psi$ has $n-2$ symbols, so by hypothesis it has an even number of parentheses.

(6) Putting the outer pair back adds 2, so χ has an even number of parentheses.

(7) By Theorem 4.2.2, every formula has an even number of parentheses.

(a) Is the conclusion true?

(b) Which line fails, and why?

(c) Show that the move made in that line is illegal by applying the same move to a different property that every formula has, and reaching a false statement.

(d) Repair the argument.

Solution. *(a)* Yes. By Lemma 4.3.2 (a), a formula has as many left as right parentheses, say k of each, so $2k$ in all.

(b) Line 5. The induction hypothesis of line 2 is about **formulas** with fewer than n symbols, and line 5 applies it to the string $\varphi \circ \psi$, which is not a formula: for $\chi = (p \wedge q)$, it is $p \wedge q$, which fails the first-symbol test of Section 4.1 (length 3, beginning with an atom). The hypothesis says nothing about such strings. Every other line is correct: line 4 applies the hypothesis to the formula φ, and lines 1, 2, 3, 6 and 7 are a correct frame for a strong induction on the length.

(c) Take the property of Lemma 4.3.2 (c), that no binary connective lies outside all paren-

theses; every formula has it. Run the same argument for it: in the case $\chi = (\varphi \circ \psi)$, line 5's move applies the hypothesis to the string $\varphi \circ \psi$ and asserts that no binary connective of $\varphi \circ \psi$ lies outside all parentheses. For $\chi = (p \wedge q)$ this asserts it of $p \wedge q$, whose $\wedge$ lies outside all parentheses. So the move leads from true hypotheses to a false statement; it is not a valid step, even though in the original argument its conclusion happens to be true.

(d) In line 5, apply the hypothesis to the formulas φ and ψ, each with fewer than n symbols: they have $2a$ and $2b$ parentheses for some integers $a, b \geq 0$, and χ has $2a + 2b + 2$, which is even. (Alternatively, prove the claim by Theorem 4.3.1, with the same three cases.)

Marking scheme. (10 points)

- 2: (a), with Lemma 4.3.2 (a).
- 3: (b), naming line 5 and why it fails (the hypothesis is about formulas, and $\varphi \circ \psi$ is not one).
- 3: (c), a property every formula has and a formula on which the move leads to a false statement.
- 2: (d), the repair through the formulas φ and ψ.

Fatal errors (0 for the step): locating the error in line 4.

Exercise. C3: Exit check: structural induction

Say that a formula is a **$\neg\wedge$-formula** if its only connectives are $\neg$ and $\wedge$ and it does not contain $\bot$. Define, by Theorem 4.3.4, the number of nodes of the parse tree and the number of negations:

$$\begin{aligned} \mathrm{nodes}(p) &= 1, \\ \mathrm{nodes}(\neg\varphi) &= \mathrm{nodes}(\varphi) + 1, \\ \mathrm{nodes}((\varphi \wedge \psi)) &= \mathrm{nodes}(\varphi) + \mathrm{nodes}(\psi) + 1, \\ \mathrm{neg}(p) &= 0, \\ \mathrm{neg}(\neg\varphi) &= \mathrm{neg}(\varphi) + 1, \\ \mathrm{neg}((\varphi \wedge \psi)) &= \mathrm{neg}(\varphi) + \mathrm{neg}(\psi), \end{aligned}$$

with $\mathrm{nodes}(\bot) = 1$ and $\mathrm{neg}(\bot) = 0$, and the clauses for the other binary connectives as for $\wedge$ (these choices play no part in what follows). Prove by structural induction that every $\neg\wedge$-formula φ satisfies $\mathrm{nodes}(\varphi) > \mathrm{neg}(\varphi)$.

Solution (Hint 1). Theorem 4.3.1 is about all formulas. Put the restriction to $\neg\wedge$-formulas inside the property.

Solution (Hint 2). Let $P(\varphi)$ be "if φ is a $\neg\wedge$-formula, then $\mathrm{nodes}(\varphi) > \mathrm{neg}(\varphi)$". The immediate parts of a $\neg\wedge$-formula are $\neg\wedge$-formulas.

Solution. For a formula φ, let $P(\varphi)$ be: "if φ is a $\neg\wedge$-formula, then $\mathrm{nodes}(\varphi) > \mathrm{neg}(\varphi)$". We prove $P(\varphi)$ for every formula φ by Theorem 4.3.1; for a $\neg\wedge$-formula it is the claim.

The immediate parts of a $\neg\wedge$-formula are $\neg\wedge$-formulas, since every symbol of a part is a

symbol of the whole.

Atoms. $\text{nodes}(p) = 1 > 0 = \text{neg}(p)$.

$\bot$, and the connectives $\vee, \to, \leftrightarrow$. The formulas $\bot$ and $(\varphi \circ \psi)$ with $\circ$ one of $\vee, \to, \leftrightarrow$ are not $\neg\wedge$-formulas, so for them P holds vacuously, whatever the clauses chosen for them.

Negation. Suppose $P(\varphi)$, and let $\neg\varphi$ be a $\neg\wedge$-formula. Then φ is one, so $\text{nodes}(\varphi) > \text{neg}(\varphi)$ by $P(\varphi)$, and adding 1 to both sides gives $\text{nodes}(\neg\varphi) > \text{neg}(\neg\varphi)$.

Conjunction. Suppose $P(\varphi)$ and $P(\psi)$, and let $(\varphi \wedge \psi)$ be a $\neg\wedge$-formula. Then φ and ψ are, so $\text{nodes}(\varphi) > \text{neg}(\varphi)$ and $\text{nodes}(\psi) > \text{neg}(\psi)$. Adding the two inequalities and then 1,

$$\text{nodes}((\varphi \wedge \psi)) = \text{nodes}(\varphi) + \text{nodes}(\psi) + 1 > \text{neg}(\varphi) + \text{neg}(\psi) = \text{neg}((\varphi \wedge \psi)).$$

(Since the counts are integers, each hypothesis gives at least 1 to spare, so the left side exceeds the right by at least 3.)

By Theorem 4.3.1, $P(\varphi)$ holds for every formula, which proves the claim.

Marking scheme. (10 points)

- 2: states the property for every formula, in a form that covers exactly the $\neg\wedge$-formulas, with the recursive counts.
- 1: atom case.
- 2: negation case, from the hypothesis for φ, after checking that φ is a $\neg\wedge$-formula.
- 3: conjunction case, using the hypotheses for **both** parts.
- 2: names the induction principle and says why $\bot$ and the other connectives are harmless (vacuous).

Fatal errors (0 for the step): applying the hypothesis to the formula being proved; a conjunction case that uses one part only and says "similarly" for the other.

Exercise. C4: Challenge: Inside the folded proofs

These parts are about the proofs in the second-reading folds of this section; read the folds first.

(a) In the proof of Theorem 4.3.1, on what is the strong induction performed?

(b) Where in the proof of Theorem 4.3.4 is Theorem 4.3.3 used, and why is it not needed for uniqueness?

(c) In the fold "Sets built from a base", prove part (i).

Solution. (a) On the position i of an entry in a formation sequence $\chi_1, \ldots, \chi_m$ of the given formula, with the property "if $1 \le i \le m$, then χ_i has P". Not on the height, which at that point is not yet known to be well defined.

(b) In the inductive step, to define $h(\chi)$ for a formula χ of length $n+1$: unique readability says which clause applies and to which parts, so the clauses give exactly one value. That two functions obeying the clauses agree needs no unique readability: by Theorem 4.3.1, if f and f' both obey the clauses, they agree at atoms and $\bot$; and if they agree at φ (and ψ), then at $\neg\varphi$ both equal $g_\neg(\varphi, f(\varphi))$, and at $(\varphi \circ \psi)$ both equal $g_\circ(\varphi, \psi, f(\varphi), f(\psi))$, whichever way

the string may be split.

(c) Let C be a closed collection, χ a formula, and $\chi_1, \dots, \chi_m$ a formation sequence with $\chi_m = \chi$. We prove, by Theorem 4.2.2 on i, the property "if $1 \le i \le m$, then $\chi_i \in C$". Let $i \in \mathbb{N}$ and suppose the property for every $k < i$. If $i = 0$ or $i > m$, it holds vacuously. Otherwise: an atom or $\bot$ is in C; if $\chi_i = \neg\chi_j$ with $j < i$, then $\chi_j \in C$ by hypothesis, so $\chi_i \in C$; if $\chi_i = (\chi_j \circ \chi_l)$ with $j, l < i$, then $\chi_j, \chi_l \in C$, so $\chi_i \in C$. By Theorem 4.2.2 the property holds for every i, and at $i = m$ it gives $\chi \in C$. Since every formula is the last entry of some formation sequence, every formula is in C.

Marking scheme. (6 points)

- 2: (a).
- 2: (b).
- 2: (c): 1 for the strong induction set up on positions, 1 for the three cases.

Exercise. C5: Challenge: formulas without parentheses

Write connectives in front of their parts and drop all parentheses. The **Polish formulas** are given by the clauses: every atom and $\bot$ is a Polish formula; if φ is one, so is $\mathsf{N}\varphi$; if φ and ψ are, so are $\mathsf{K}\varphi\psi$, $\mathsf{A}\varphi\psi$, $\mathsf{C}\varphi\psi$ and $\mathsf{E}\varphi\psi$ (for $\wedge, \vee, \to, \leftrightarrow$); and nothing else is. Give each symbol a **weight**: $+1$ for an atom or $\bot$, 0 for N, -1 for each of $\mathsf{K}, \mathsf{A}, \mathsf{C}, \mathsf{E}$. The weight $w(u)$ of a string u is the sum of the weights of its symbols.

(a) Write $((p \wedge \neg q) \to p)$ and $\neg(p \vee (q \leftrightarrow r))$ as Polish formulas.

(b) Prove that every Polish formula has weight 1, and that every proper initial segment of it (the empty one included) has weight at most 0.

(c) Deduce that if $\mathsf{K}\varphi\psi$ and $\mathsf{K}\varphi'\psi'$ are the same string, for Polish formulas $\varphi, \psi, \varphi', \psi'$, then $\varphi = \varphi'$ and $\psi = \psi'$.

(d) Read $\mathsf{CA}pq\mathsf{N}r$ by the count, and write it in ordinary notation.

Solution (Hint 1).

(b) is an induction over the Polish clauses, set up as in the fold "Sets built from a base": the Polish formulas with the property form a collection closed under the Polish clauses. Use a strengthened property, as in Lemma 4.3.2.

Solution (Hint 2). (c): if $\varphi \neq \varphi'$, one of them is a proper initial segment of the other.

Solution. (a) $\mathsf{CK}p\mathsf{N}qp$ and $\mathsf{NA}p\mathsf{E}qr$.

Read "nothing else is" as in the fold "Sets built from a base": the Polish formulas are the smallest collection of strings that contains every atom and $\bot$ and is closed under the Polish clauses. So if a property holds at atoms and $\bot$ and passes from the parts to the whole at each clause, the Polish formulas with it form such a collection and include every Polish formula. In the same way every Polish formula is non-empty, and is an atom, $\bot$, $\mathsf{N}\varphi$, or $\mathsf{K}\varphi\psi$, $\mathsf{A}\varphi\psi$, $\mathsf{C}\varphi\psi$, $\mathsf{E}\varphi\psi$ for Polish φ, ψ (the strings of these forms make such a collection); (c) and (d) use

this.

(b) By induction over the clauses, the property being "weight 1, and every proper initial segment has weight ≤ 0".

Atom or $\bot$. Weight 1; its only proper initial segment is the empty string, of weight 0.

$\mathsf{N}\varphi$. Weight $0 + 1 = 1$. A proper initial segment is empty (weight 0), or N followed by a proper initial segment u of φ (weight $0 + w(u) \leq 0$).

$\mathsf{K}\varphi\psi$ (and likewise $\mathsf{A}, \mathsf{C}, \mathsf{E}$). Weight $-1+1+1 = 1$. A proper initial segment is of one of four kinds: empty (weight 0); K followed by an initial segment u of φ other than φ itself (weight $-1 + w(u) \leq -1$); $\mathsf{K}\varphi$ (weight $-1 + 1 = 0$); or $\mathsf{K}\varphi$ followed by a proper non-empty initial segment u of ψ (weight $0 + w(u) \leq 0$).

(c) Both φ and φ' begin at the second symbol. If they differ, the shorter is an initial segment of the longer (both are initial segments of the string after K), a proper one, and non-empty, since every Polish formula has a symbol. By (b) the shorter has weight 1, and as a proper initial segment of the longer it has weight ≤ 0: impossible. So $\varphi = \varphi'$, and ψ, ψ' are both what follows $\mathsf{K}\varphi$: equal.

(d) Read from the symbol after C: the running weight is -1 after A, 0 after p, and first reaches 1 after q. By (b), a Polish formula ends exactly where its running weight first reaches 1, so the first part of C is $\mathsf{A}pq$ and the second is $\mathsf{N}r$. (From the first symbol the running weight is $-1, -2, -1, 0, 0, 1$, reaching 1 only at the end.) In ordinary notation: $((p \vee q) \to \neg r)$.

Marking scheme. (10 points)

- 2: (a).
- 4: (b), with the induction set up over the clauses and all three kinds of non-empty proper initial segment of $\mathsf{K}\varphi\psi$.
- 3: (c), with the non-emptiness of the shorter part named.
- 1: (d).

Fatal errors (0 for the step): in (b), proving only "weight 1" (then (c) has nothing to use: the strengthening is the point); in (c), assuming the two parts have equal length.

4.4 Valuations and Truth Tables

Sections 4.1 and 4.3 made formulas into objects: every formula is built from atoms and $\bot$ by the connectives in exactly one way, and so a function on formulas can be defined one clause per connective. So far a formula is only a string of symbols. It says nothing, and it is neither true nor false. This section gives formulas a meaning. We say what it takes to fix the truth or falsity of every atom, and then compute the truth value of every formula from the values of its atoms, one connective at a time. The result is the truth table, the second way of looking at a formula that Chapters 1 and 2 kept promising.

Throughout, formulas are those of Definition 4.1.1, with atoms $p, q, r, p_1, p_2, \ldots$, and we use Chapter 1's shorthand for them (outer parentheses dropped, the usual precedence, $\to$ grouped to the right), as Section 4.1 allows.

4.4.1 Reading a table off the rules

Before any definition, a question. Suppose we want to call each statement true or false, and we want the rules of Chapters 1 and 2 to be **safe**: whenever the lines a rule cites are true, the line it writes is true. How must "true" behave for each connective? The rules answer almost completely. This subsection is motivation only; nothing later rests on it, and Section 4.8 makes it exact.

Conjunction. Take statements φ and ψ. If both are true, $\wedge$I writes $\varphi \wedge \psi$ from them, so $\varphi \wedge \psi$ must be true. If $\varphi \wedge \psi$ is true, $\wedge$E writes φ and writes ψ, so both must be true; turned around, if either of them is false, then $\varphi \wedge \psi$ must be false. All four combinations of values are settled:

φ	ψ	$\varphi \wedge \psi$
T	T	T
T	F	F
F	T	F
F	F	F

Here T stands for "true" and F for "false".

Contradiction. The rule $\bot$E writes **any** statement from $\bot$. If $\bot$ were ever true, $\bot$E would lead from a true line to every statement, false ones included. So $\bot$ must be false, always.

Negation. The rule $\neg$E writes $\bot$ from φ and $\neg\varphi$. Since $\bot$ is false, φ and $\neg\varphi$ are never both true: if φ is true, $\neg\varphi$ is false. And Proposition 2.5.3 derives $\varphi \vee \neg\varphi$ from no premises at all, so at least one of the two is true: if φ is false, $\neg\varphi$ is true. Negation flips the value.

Implication. The rule $\to$E writes ψ from $\varphi \to \psi$ and φ. So in the combination "φ true, ψ false", the implication cannot be true: that is the one row where $\varphi \to \psi$ is false. The other three rows are forced by derivations. Proposition 2.1.2 derives $P \to Q$ from $\neg P$, so whenever P is false (and $\neg P$ therefore true), $P \to Q$ must be true, whatever Q is. And $P \to Q$ is derivable from Q alone (Exercise (B2: An assumption that is not used) (a) of Section 1.4):

1	Q	
2	$\quad P$	asm, for $\to$I
3	$\quad Q$	R 1
4	$P \to Q$	$\to$I 2–3

so whenever Q is true, $P \to Q$ must be true. That settles the rows with P false (by the first derivation) and the row with both true (by the second).

The vacuous rows of $\to$, "if P is false, then $P \to Q$ is true", are often presented as a convention to be swallowed. Here they are forced by $\bot$E, the rule that closes an impossible case, exactly as Section 2.1 said they would be. The rows of $\vee$ and $\leftrightarrow$ can be read off their rules in the same way, and Exercise C3 of Section 4.5 asks for a sharper version of the argument for $\to$. We now turn the answers into a definition.

Second reading · *A pair of rules with no table*

The rules of Chapters 1 and 2 determine their tables almost completely. Not every pair of rules does. Invent a connective $\diamond$ with an introduction rule "from φ, infer $\varphi \diamond \psi$" and an elimination

rule "from $\varphi \diamond \psi$, infer ψ", and read a table off them as above. The first rule needs $\varphi \diamond \psi$ true whenever φ is; the second needs ψ true whenever $\varphi \diamond \psi$ is. In a row with φ true and ψ false the two demands clash, so no table makes both rules safe. Worse, the two rules together derive anything from anything: from p, the first gives $p \diamond q$, and the second then gives q. (Logicians call this connective *tonk*.) A pair of rules does not give a connective a meaning just by being written down. Section 4.8 shows that the rules of Chapters 1 and 2 do, row by row.

4.4.2 Valuations

A formula such as $p \wedge \neg q$ cannot be true or false on its own, because p and q are atoms: letters with no content. What decides its value is a choice of which atoms are true. We need a word for one such choice.

A valuation is one possible situation, described by saying which atoms are true and which are false.

Definition 4.4.1. Valuation

A **valuation** v assigns to **every** atom p **exactly one** of the two truth values T and F. We write $v(p)$ for the value it assigns to p, and say that p is **true under** v if $v(p) = \mathrm{T}$, **false under** v if $v(p) = \mathrm{F}$.

In words: a valuation answers, for each atom separately, the question "true or false?", and it answers every such question. Two valuations are equal when they give every atom the same value. Nothing links the values of different atoms: $v(p)$ and $v(q)$ may be chosen freely, and that freedom is the whole point.

We let a valuation decide **every** atom so that one valuation gives a value to every formula at once. In practice only the few atoms a formula contains matter, and the Coincidence Lemma below says so.

Some valuations, simplest first.

- The **all-T valuation** gives every atom the value T; the **all-F valuation** gives every atom F. They are the two extreme cases, and both will return as tests.
- The valuation with $v(p) = \mathrm{T}$, $v(q) = \mathrm{F}$ and $v(r) = \mathrm{F}$ for every other atom r. This is the typical way to describe a valuation: give the values of the few atoms we care about, and fix the rest once and for all.

4.4.3 The value of a formula

A valuation fixes the atoms. The value of every other formula should then follow, computed from the values of its parts by the tables of the first subsection. Since formulas are built by the clauses of Definition 4.1.1, this is a definition by recursion.

The value of a formula is computed from the bottom of its parse tree upwards, one connective at a time.

Definition 4.4.2. Truth Value of a Formula

Let v be a valuation. The **truth value** $v(\varphi)$ of a formula φ **under** v is defined by recursion

on φ:

(V1) for an atom p, $v(p)$ is the value the valuation assigns to p; and $v(\bot) \coloneqq \mathrm{F}$;
(V2) $v(\neg\varphi) \coloneqq \mathrm{T}$ if $v(\varphi) = \mathrm{F}$, and $v(\neg\varphi) \coloneqq \mathrm{F}$ if $v(\varphi) = \mathrm{T}$;
(V3) $v(\varphi \wedge \psi) \coloneqq \mathrm{T}$ if $v(\varphi) = \mathrm{T}$ **and** $v(\psi) = \mathrm{T}$; otherwise F;
(V4) $v(\varphi \vee \psi) \coloneqq \mathrm{T}$ if $v(\varphi) = \mathrm{T}$ **or** $v(\psi) = \mathrm{T}$ (or both); otherwise F;
(V5) $v(\varphi \to \psi) \coloneqq \mathrm{F}$ if $v(\varphi) = \mathrm{T}$ and $v(\psi) = \mathrm{F}$; otherwise T;
(V6) $v(\varphi \leftrightarrow \psi) \coloneqq \mathrm{T}$ if $v(\varphi) = v(\psi)$; otherwise F.

We say φ is **true under** v, or v **makes** φ true, if $v(\varphi) = \mathrm{T}$; otherwise φ is **false under** v.

Clauses (V2)–(V6) say that the value of a compound formula depends **only** on the values of its immediate parts, and they are summed up in the tables of the five connectives:

φ	$\neg\varphi$
T	F
F	T

φ	ψ	$\varphi \wedge \psi$	$\varphi \vee \psi$	$\varphi \to \psi$	$\varphi \leftrightarrow \psi$
T	T	T	T	T	T
T	F	F	T	F	F
F	T	F	T	T	F
F	F	F	F	T	T

Read them clause by clause. (V1): an atom has the value the valuation gives it, and $\bot$ is false under every valuation. (V2): negation flips. (V3): a conjunction needs both parts. (V4): a disjunction needs at least one part; "or" is **inclusive**, as the rule $\vee$I of Section 1.5 already said, so the row with both parts true gives T. (V5): an implication is false in exactly one case, a true hypothesis with a false conclusion. (V6): a biconditional says the two sides agree.

The columns for $\neg$, $\wedge$ and $\to$, and the value of $\bot$, are the ones the first subsection read off the rules. Section 4.8 proves that every row of every table is, in a precise sense, a derivable sequent.

Well defined. Clause (V1) gives a value to every atom and to $\bot$, and each of (V2)–(V6) gives a rule computing the value of a compound formula from the values of its immediate parts. By Theorem 4.3.4, such clauses define exactly one function on all formulas. This is what Section 4.3 was for. Section 4.3 showed the danger with the string $p \wedge q \vee r$, whose two readings take different values when p and q are false and r is true. A formula of Definition 4.1.1 carries its parentheses, and has one value.

Example. Computing a value

Let v be a valuation with $v(p) = \mathrm{F}$, $v(q) = \mathrm{T}$ and $v(r) = \mathrm{F}$. Compute $v(\varphi)$ for

$$\varphi = (p \vee \neg q) \to (q \wedge r).$$

Solution. Work from the leaves of the parse tree up to its root, one node at a time, each time using one clause of Definition 4.4.2. List the subformulas so that each comes after its immediate parts:

Subformula	Clause	Value
p, q, r	(V1)	F, T, F
$\neg q$	(V2): $v(q) = \mathrm{T}$	F

$p \vee \neg q$	(V4): both parts F	F
$q \wedge r$	(V3): $v(r) = \mathrm{F}$	F
$(p \vee \neg q) \to (q \wedge r)$	(V5): the hypothesis is F	T

So $v(\varphi) = \mathrm{T}$. The last step did not need the value of $q \wedge r$: a false hypothesis makes an implication true whatever its conclusion is. Computing by hand, one may stop a branch as soon as its value is decided.

↻ **Second reading** · *Edge cases*

Nonexample. "p is true" and "p is false" together do not describe a valuation: the definition asks for **exactly one** value per atom. And a list that gives p the value T and says nothing about the other atoms is not a valuation either, since it fails "every atom". We never need more than such a partial list, but we complete it (by F, say) so that the definition applies.

Nonexample. The rule "every formula gets the value T" agrees with the all-T valuation on every atom, but it is not the truth value of Definition 4.4.2 under any valuation. It breaks (V1) at $\bot$, and it breaks (V2) at $\neg p$, where p and $\neg p$ would both be T. The truth value under the all-T valuation is the function that (V1)–(V6) produce from it, and it makes $\neg p$ false.

Warning. Values do not distribute over connectives the way symbols seem to. The value of $\neg(p \wedge q)$ is not computed by negating p and q and keeping the $\wedge$: under a valuation with $v(p) = \mathrm{T}$ and $v(q) = \mathrm{F}$, the formula $\neg(p \wedge q)$ is T (since $p \wedge q$ is F), while $\neg p \wedge \neg q$ is F (since $\neg p$ is F). Always follow the parse tree: the main connective is applied last.

Quick check. Let $v(p) = \mathrm{F}$, $v(q) = \mathrm{T}$, $v(r) = \mathrm{T}$. What is $v\big((p \vee q) \to \neg r\big)$?

Solution. F. Bottom up: $p \vee q$ is T, by (V4), since q is T; $\neg r$ is F, by (V2); and an implication with a true hypothesis and a false conclusion is F, by (V5).

Quick check. Let $v(p) = \mathrm{T}$ and $v(q) = \mathrm{F}$. What is $v(\neg p \vee (q \leftrightarrow \bot))$?

Solution. T. $\neg p$ is F by (V2). $q \leftrightarrow \bot$ is T by (V6), since both sides are F ($\bot$ by (V1)). So the disjunction is T by (V4).

A game for the clauses. Clauses (V2)–(V4) can be acted out. Two players argue about a formula built from atoms by $\neg$, $\wedge$ and $\vee$, under a valuation v. The **verifier** claims that the formula is true under v, the **falsifier** that it is false. Play starts at the whole formula and moves down its parse tree. At $\varphi \wedge \psi$ the falsifier chooses a conjunct to dispute, and play goes on there;

at $\varphi \vee \psi$ the verifier chooses a disjunct to defend. At $\neg\varphi$ the two players swap roles, and play goes on at φ. At an atom p the game ends: the verifier wins if $v(p) = \mathrm{T}$, the falsifier if $v(p) = \mathrm{F}$. Run it once on $\neg(p \wedge q)$, with $v(p) = \mathrm{T}$ and $v(q) = \mathrm{F}$. At the $\neg$ the players swap, so the verifier of $\neg(p \wedge q)$ now disputes $p \wedge q$. At the $\wedge$ the choice is hers: she picks q, and she wins. A formula is true under v exactly when the verifier can always win, whatever the falsifier does; Exercise C3 proves it, and nothing later rests on the game.

4.4.4 Only the atoms of a formula matter

A valuation fixes infinitely many values, but a formula contains only finitely many atoms. We expect the value of a formula to depend only on the values of its own atoms. That is the fact that makes truth tables finite, and it is our first proof by structural induction about meaning.

Recall from Section 4.3 that $\mathrm{atoms}(\varphi)$, the set of atoms occurring in φ, is defined by recursion: $\mathrm{atoms}(p) = \{p\}$ for an atom p, $\mathrm{atoms}(\bot)$ is empty, $\mathrm{atoms}(\neg\varphi) = \mathrm{atoms}(\varphi)$, and the atoms of $\varphi \circ \psi$, for each binary connective $\circ$, are those of φ together with those of ψ. Two valuations v and w **agree on** a set of atoms if $v(p) = w(p)$ for every atom p in it.

Lemma 4.4.1. Coincidence Lemma

Let φ be a formula, and let v and w be valuations that agree on $\mathrm{atoms}(\varphi)$. Then $v(\varphi) = w(\varphi)$.

Idea. Structural induction on φ, with the property "for **all** valuations v, w that agree on the atoms of φ, $v(\varphi) = w(\varphi)$". At an atom, the two valuations agree by hypothesis. At a compound formula, both values are computed by the same clause from the values of the immediate parts, and every atom of a part is an atom of the whole, so the two valuations agree on the atoms of each part and the hypothesis for the parts applies.

Proof. We apply Theorem 4.3.1 to the property $C(\varphi)$: *for all valuations v and w that agree on* $\mathrm{atoms}(\varphi)$, $v(\varphi) = w(\varphi)$.

Atoms. Let p be an atom and let v, w agree on $\mathrm{atoms}(p) = \{p\}$. Then $v(p) = w(p)$, which is $C(p)$.

$\bot$. By (V1), $v(\bot) = \mathrm{F} = w(\bot)$ for all valuations v, w. So $C(\bot)$ holds.

Negation. Assume $C(\varphi)$, and let v, w agree on $\mathrm{atoms}(\neg\varphi)$. Since $\mathrm{atoms}(\neg\varphi) = \mathrm{atoms}(\varphi)$, they agree on $\mathrm{atoms}(\varphi)$, so $v(\varphi) = w(\varphi)$ by $C(\varphi)$. By (V2), $v(\neg\varphi)$ is determined by $v(\varphi)$ and $w(\neg\varphi)$ by $w(\varphi)$ through the same rule, so $v(\neg\varphi) = w(\neg\varphi)$. This proves $C(\neg\varphi)$.

Conjunction. Assume $C(\varphi)$ and $C(\psi)$, and let v, w agree on $\mathrm{atoms}(\varphi \wedge \psi)$. Every atom of φ is an atom of $\varphi \wedge \psi$, so v and w agree on $\mathrm{atoms}(\varphi)$, and $v(\varphi) = w(\varphi)$ by $C(\varphi)$. In the same way $v(\psi) = w(\psi)$ by $C(\psi)$. By (V3), $v(\varphi \wedge \psi) = \mathrm{T}$ exactly when $v(\varphi) = v(\psi) = \mathrm{T}$, that is, exactly when $w(\varphi) = w(\psi) = \mathrm{T}$, that is, exactly when $w(\varphi \wedge \psi) = \mathrm{T}$. Hence $v(\varphi \wedge \psi) = w(\varphi \wedge \psi)$, which is $C(\varphi \wedge \psi)$.

Disjunction, implication, biconditional. Assume $C(\varphi)$ and $C(\psi)$, and let v, w agree on the atoms of $\varphi \circ \psi$. As in the conjunction case, $v(\varphi) = w(\varphi)$ and $v(\psi) = w(\psi)$. One clause of Definition 4.4.2 then turns these equal inputs into equal outputs:

$\circ$	Clause	Why $v(\varphi \circ \psi) = w(\varphi \circ \psi)$
$\vee$	(V4)	T exactly when one of the two parts is T, and the parts have the same values under v and under w
$\to$	(V5)	F exactly when the first part is T and the second F, and the parts have the same values under v and under w
$\leftrightarrow$	(V6)	T exactly when the two parts have equal values, and the parts have the same values under v and under w

Every clause of Definition 4.1.1 has been covered, so by Theorem 4.3.1, $C(\varphi)$ holds for every formula φ. This proves the lemma.

■

The table's three cases are the conjunction case again, with another clause. The fold writes each one out in words.

↻ **Second reading** · *The three remaining cases, written out*

Disjunction. Assume $C(\varphi)$ and $C(\psi)$, and let v, w agree on atoms$(\varphi \vee \psi)$. As in the previous case, $v(\varphi) = w(\varphi)$ and $v(\psi) = w(\psi)$. By (V4), $v(\varphi \vee \psi) = \mathrm{T}$ exactly when one of $v(\varphi)$, $v(\psi)$ is T, that is, when one of $w(\varphi)$, $w(\psi)$ is T, that is, when $w(\varphi \vee \psi) = \mathrm{T}$. So $C(\varphi \vee \psi)$.

Implication. Assume $C(\varphi)$ and $C(\psi)$, and let v, w agree on atoms$(\varphi \to \psi)$. Again $v(\varphi) = w(\varphi)$ and $v(\psi) = w(\psi)$. By (V5), $v(\varphi \to \psi) = \mathrm{F}$ exactly when $v(\varphi) = \mathrm{T}$ and $v(\psi) = \mathrm{F}$, that is, when $w(\varphi) = \mathrm{T}$ and $w(\psi) = \mathrm{F}$, that is, when $w(\varphi \to \psi) = \mathrm{F}$. So $C(\varphi \to \psi)$.

Biconditional. Assume $C(\varphi)$ and $C(\psi)$, and let v, w agree on atoms$(\varphi \leftrightarrow \psi)$. Again $v(\varphi) = w(\varphi)$ and $v(\psi) = w(\psi)$. By (V6), $v(\varphi \leftrightarrow \psi) = \mathrm{T}$ exactly when $v(\varphi) = v(\psi)$, that is, when $w(\varphi) = w(\psi)$, that is, when $w(\varphi \leftrightarrow \psi) = \mathrm{T}$. So $C(\varphi \leftrightarrow \psi)$.

Two consequences. First, to know $v(\varphi)$ we need only the values of v on the finitely many atoms of φ: if φ has n atoms, it has exactly 2^n different "situations", one for each way of giving its atoms the values T and F. (Each atom has two choices, and the choices are independent, so the number of ways doubles with each new atom: $2, 4, 8, \ldots, 2^n$. This count is itself a small induction on n, by Axiom 4.2.1.) Second, the extra atoms a valuation happens to fix are irrelevant: two valuations that differ only on atoms not in φ give φ the same value.

Quick check. Let $\varphi = p \to (q \vee p)$. Two valuations v and w agree on p and q but disagree on r and on every p_n. What does Lemma 4.4.1 say about $v(\varphi)$ and $w(\varphi)$? Does it say anything about $v(\varphi \wedge r)$ and $w(\varphi \wedge r)$?

Solution. $\text{atoms}(\varphi) = \{p, q\}$, and v, w agree there, so $v(\varphi) = w(\varphi)$. For $\varphi \wedge r$ the atoms are p, q, r, and the valuations disagree on r, so the lemma does not apply. Indeed the values can differ: φ is true under every valuation (if p is true, so is $q \vee p$), so $v(\varphi \wedge r) = v(r)$ and $w(\varphi \wedge r) = w(r)$, which differ.

4.4.5 Truth tables

Lemma 4.4.1 says that a formula's value is decided by finitely many choices, and so it can be tabulated once and for all.

A truth table lists the value of a formula in every possible situation, one row per way of giving its atoms values.

Definition 4.4.3. Truth Table

Let φ be a formula and let $p_1, \ldots, p_n$ be a list of distinct atoms that **includes every atom of** φ. A **row** is an assignment of a truth value to each of $p_1, \ldots, p_n$. The **truth table** of φ (on $p_1, \ldots, p_n$) lists, for each of the 2^n rows, the value $v(\varphi)$ for a valuation v that **agrees with the row** on $p_1, \ldots, p_n$. We say φ is true (or false) **in** the row.

Well defined. Two things must be checked. **Existence:** every row comes from a valuation, for instance the one that gives $p_1, \ldots, p_n$ the row's values and every other atom F. **Independence of the choice:** two valuations that agree with the same row agree on $p_1, \ldots, p_n$, hence on $\text{atoms}(\varphi)$, so by Lemma 4.4.1 they give φ the same value. Without the lemma, "the value in a row" would not make sense, since a row says nothing about the infinitely many other atoms.

The list $p_1, \ldots, p_n$ is allowed to contain atoms that do not occur in φ; then pairs of rows that differ only there show the same value. This freedom lets us compare several formulas in one table, over the atoms of all of them, and Section 4.5 does that constantly. When φ has no atoms at all, as $\neg\bot$ or $\bot \to \bot$ do, the list may be empty, and the table has $2^0 = 1$ row: such a formula has the same value under every valuation.

Conventions. The atom columns come first. The rows are listed with T before F, the last atom alternating fastest: for two atoms, TT, TF, FT, FF. In practice we add a column for each subformula, building up to φ in the order of the parse tree, as in Example (Computing a value), but now for every row at once.

Example. The truth table of modus ponens

Make the truth table of $((p \to q) \wedge p) \to q$ on p, q, and read off from it what $\to$E does.

Solution. Columns for the two atoms, then for each subformula in turn:

p	q	$p \to q$	$(p \to q) \wedge p$	$((p \to q) \wedge p) \to q$
T	T	T	T	T
T	F	F	F	T
F	T	T	F	T
F	F	T	F	T

The third column is (V5): F only in the row TF. The fourth is (V3) applied to the third column and the p column: T only in the first row. The last is (V5) again, with the fourth column as hypothesis and q as conclusion: the hypothesis is false in rows 2–4, so those rows are T, and in row 1 the conclusion q is T.

So the formula is true in every row. Read as a statement about →E: the only row in which both given lines $p \to q$ and p are true is row 1, and in that row the line the rule writes, q, is true as well. The rule never leads from true lines to a false one, at least when φ and ψ are the atoms p and q. Section 4.7 proves this for every rule and every formula.

Now a table built column by column for a biconditional, whose two sides Chapter 2 related by derivations.

Example. An implication and its contrapositive

Make the truth table of $(p \to q) \leftrightarrow (\neg q \to \neg p)$.

Solution.

p	q	$p \to q$	$\neg q$	$\neg p$	$\neg q \to \neg p$	$(p \to q) \leftrightarrow (\neg q \to \neg p)$
T	T	T	F	F	T	T
T	F	F	T	F	F	T
F	T	T	F	T	T	T
F	F	T	T	T	T	T

The columns $p \to q$ and $\neg q \to \neg p$ are identical: each is F only in row 2. For $\neg q \to \neg p$ that is the row where the hypothesis $\neg q$ is T and the conclusion $\neg p$ is F, that is, q false and p true, which is the same row TF. So the biconditional is T in every row.

Compare Chapter 2. There Proposition 2.3.2 derived $\neg Q \to \neg P$ from $P \to Q$ with the rules for implication and negation, and Proposition 2.5.2 derived the way back with a derivation that uses RAA. The table sees no difference between the two directions; it only sees that the two columns agree. What a table can and cannot tell about derivations is the subject of Sections 4.7 and 4.8.

A table with n atoms has 2^n rows, so three atoms give 8 and five give 32. Often we do not need all of them.

Example. Working backwards from F

Find every row in which

$$\varphi = (p \to q) \to (p \to (q \wedge r))$$

is false: first with the full table on p, q, r, then without it.

Solution. The full table.

p	q	r	$p \to q$	$q \wedge r$	$p \to (q \wedge r)$	$(p \to q) \to (p \to (q \wedge r))$
T	T	T	T	T	T	T
T	T	F	T	F	F	F
T	F	T	F	F	F	T
T	F	F	F	F	F	T
F	T	T	T	T	T	T
F	T	F	T	F	T	T
F	F	T	T	F	T	T
F	F	F	T	F	T	T

The formula is false in exactly one row, row 2: p and q true, r false.

Backwards from F. An implication is false in exactly one way, by (V5): hypothesis true, conclusion false. Apply that to the main connective of φ, and then to whatever it forces.

1. φ is F exactly when $p \to q$ is T and $p \to (q \wedge r)$ is F.
2. $p \to (q \wedge r)$ is F exactly when p is T and $q \wedge r$ is F.
3. With p true, $p \to q$ is T exactly when q is T.
4. With q true, $q \wedge r$ is F exactly when r is F.

Each step was forced, so there is exactly one row making φ false: $p = \mathrm{T}$, $q = \mathrm{T}$, $r = \mathrm{F}$. Four short steps replace eight rows.

The second method has a name in this book: **working backwards from F**. Assume the formula is false, and ask what each part is forced to be, starting from the main connective. Implications and disjunctions are false in only one way, and conjunctions are true in only one way, so they force values; a false conjunction, a true disjunction, or a true implication leaves a choice, and then we split into cases. If every case ends in a contradiction (some atom would need both values), the formula is false in **no** row. A case that ends in a consistent assignment, completed in every way on the atoms it leaves unset, gives rows making the formula false, and together these cases give all such rows.

4.4.6 Truth-functional connectives

Each of the clauses (V2)–(V6) computes the value of a compound formula from the values of its immediate parts **alone**. Nothing else about the parts enters: not what they say, not how long they are. The general form of such a rule has a name.

Definition 4.4.4. Truth Function

Let $n \geq 1$. A **truth function** f with n inputs assigns to **each** list $(a_1, \ldots, a_n)$ of truth values (each a_i is T or F) a truth value $f(a_1, \ldots, a_n)$. A formula φ **expresses** f if

$$v(\varphi) = f(v(p_1), \ldots, v(p_n)) \quad \text{for every valuation } v.$$

In words: a truth function is a truth table with n input columns and one output column, with no formula attached. Each of our connectives is a truth function. Negation is the one-input function $\mathrm{T} \mapsto \mathrm{F}$, $\mathrm{F} \mapsto \mathrm{T}$; conjunction is the two-input function that is T only at (T, T);

disjunction, implication and the biconditional are the two-input functions read off the tables of (V4), (V5) and (V6). A way of joining statements whose value is fixed by the values of its parts is called **truth-functional**, and (V2)–(V6) say that all five connectives are.

Not every way of joining two statements is. Take "φ because ψ":

φ	ψ	φ because ψ
T	T	?
T	F	F
F	T	F
F	F	F

A "because" claim with a false part is false, so three rows are settled. The first is not. "6 is even because 6 is a multiple of 3" and "6 is even because 6 is a multiple of 2" both have two true parts. The first is false, since being a multiple of 3 does not make a number even; the second is true. The values of the parts do not decide the value of the whole.

The same happens with "it was proved before 1900 that φ":

φ	it was proved before 1900 that φ
T	?
F	F

A statement that was proved is true, so the second row is F. But "$\sqrt{2}$ is irrational" and "for no $n \geq 3$ does $x^n + y^n = z^n$ have a solution in positive integers" are both true, and only the first was proved before 1900. So "because" and "it was proved before 1900 that" are not connectives of our language, and no truth table can stand for them.

Quick check. Is "neither φ nor ψ" truth-functional? Give its table and a formula that expresses it.

Solution. Yes. It is T when both parts are F, and F in the other three rows, whatever the parts say. The formula $\neg p_1 \wedge \neg p_2$ expresses it.

Warning. A truth table sees only the connectives. Everything inside an atom is invisible to it. Fix a real number x. The statement "if $x > 2$, then $x^2 > 4$" is true, but its propositional form, with the atoms p: "$x > 2$" and q: "$x^2 > 4$", is $p \to q$, which is false in the row with p true and q false. That row cannot happen for any real x, and the reason lies in algebra, not in the connectives. The quantified statement "for every real x, if $x > 2$, then $x^2 > 4$" is, like "for every $\varepsilon > 0$ there is an N such that ...", a single atom for a truth table, with two rows and nothing to compute.

Quick check. Why does Definition 4.4.3 need Lemma 4.4.1? What would go wrong in a table for $p \vee q$ if the value of $p \vee q$ could depend on the atom r?

Solution. A row gives values to the listed atoms only, and many valuations agree with it, differing on the other atoms. The table records "the" value of the formula in the row, so all those valuations must give the same value; Lemma 4.4.1 guarantees it. If $v(p \vee q)$ could depend on r, the row $p = \mathrm{F}$, $q = \mathrm{F}$ would contain valuations giving different values, and there would be no single entry to write.

Remark (Looking ahead). Chapter 5 looks inside atoms.

4.4.7 Exercises

4.4.7.1 A. Check your understanding

Exercise. A1

(a) Make the truth tables of $\neg(p \wedge \neg q)$ and of $(p \vee q) \to p$.

(b) Make the truth table of $p \leftrightarrow \neg p$. Is it true in any row?

(c) Compute the value of $(p \to \bot) \vee (\bot \to p)$ under the all-F valuation, and under the all-T valuation.

(d) How many rows has the truth table of a formula on the atoms p, q, r, s?

(e) True or false, with a reason: if $v(p) \neq w(p)$, then $v(p \to (q \to p)) \neq w(p \to (q \to p))$.

Solution.

(a)

p	q	$\neg q$	$p \wedge \neg q$	$\neg(p \wedge \neg q)$
T	T	F	F	T
T	F	T	T	F
F	T	F	F	T
F	F	T	F	T

p	q	$p \vee q$	$(p \vee q) \to p$
T	T	T	T
T	F	T	T
F	T	T	F
F	F	F	T

The first formula has the same column as $p \to q$: false only in row TF.

(b) In the row $p = \mathrm{T}$, $\neg p$ is F, and in the row $p = \mathrm{F}$, $\neg p$ is T; the two sides always differ, so by (V6) the formula is F in both rows. It is true in no row.

p	$\neg p$	$p \leftrightarrow \neg p$
T	F	F
F	T	F

(c) The disjunct $\bot \to p$ has a false hypothesis, since $v(\bot) = \mathrm{F}$ under every valuation (V1), so it is T by (V5), and the disjunction is T by (V4). This holds under both valuations (and

under every other).

(d) $2^4 = 16$.

(e) False. The formula $p \to (q \to p)$ is true under every valuation: if $v(p) = \mathrm{F}$, the hypothesis is false; if $v(p) = \mathrm{T}$, the conclusion $q \to p$ has a true conclusion. So v and w give it the value T however they differ. Lemma 4.4.1 says that agreement on the atoms forces agreement on the formula; it says nothing in the other direction.

Exercise. A2: Drill

The parts run from computing one value, through partial information and backward computation, to the game and the truth-functional connectives.

(a) Let $v(p) = \mathrm{T}$ and $v(q) = \mathrm{F}$. What is $v(p \to q)$?
(b) Same v: what is $v(\neg p \to q)$?
(c) Same v: what is $v(\bot \leftrightarrow q)$?
(d) How many rows has a table on p, q, r, s, t?
(e) Only $v(p) = \mathrm{T}$ is known. Which of $p \vee q$, $p \wedge q$, $q \to p$, $\neg p \to r$ already have a value, and which value?
(f) Give a row making $(p \vee q) \to (p \wedge q)$ false.
(g) Is there a row making $p \to \neg p$ true?
(h) Work backwards from F: in which rows is $p \to (q \to r)$ false?
(i) Find every row in which $(p \vee q) \wedge \neg r$ is true, forcing what you can first.
(j) Valuations v, w agree on p and differ on q. Must $v((p \to p) \vee q) = w((p \to p) \vee q)$?
(k) In the game at $\neg(p \wedge q)$, with $v(p) = \mathrm{F}$ and $v(q) = \mathrm{T}$: who chooses at the $\wedge$, what should be chosen, and who wins?
(l) Is "not both φ and ψ" truth-functional? If so, give a formula that expresses it.

Solution.

(a) F: a true hypothesis and a false conclusion.
(b) T: the hypothesis $\neg p$ is F.
(c) T: both sides are F.
(d) $2^5 = 32$.
(e) $p \vee q$ is T (one true disjunct is enough); $q \to p$ is T (true conclusion); $\neg p \to r$ is T (false hypothesis). $p \wedge q$ waits for q.
(f) $p = \mathrm{T}$, $q = \mathrm{F}$ (or $p = \mathrm{F}$, $q = \mathrm{T}$).
(g) Yes, $p = \mathrm{F}$.
(h) Only $p = q = \mathrm{T}$, $r = \mathrm{F}$.
(i) A true conjunction forces both parts: $\neg r$ true forces $r = \mathrm{F}$. A true $p \vee q$ leaves three choices, so the rows are p, q, r = TTF, TFF, FTF.
(j) Yes, both are T, since $p \to p$ is T under every valuation. But Lemma 4.4.1 does not say so, because v and w differ on the atom q: the lemma gives a sufficient condition, not a necessary one.

(k) At the $\neg$ the players swap, so the original verifier disputes $p \wedge q$ and chooses at the $\wedge$. She chooses the conjunct p, which is false, and wins: $\neg(p \wedge q)$ is true.

(l) Yes: it is F only when both parts are T. The formula $\neg(p_1 \wedge p_2)$ expresses it.

4.4.7.2 B. Practice

Exercise. B1: Backwards from F

For each formula, work backwards from F to find every row on its atoms in which it is false, or show that there is none.

(a) $((p \vee q) \wedge ((p \to r) \wedge (q \to r))) \to r$.

(b) $(p \to (q \to r)) \to ((p \to q) \to (p \to r))$.

(c) $((p \to r) \vee (q \to r)) \to ((p \vee q) \to r)$.

(d) $((p \to q) \wedge (q \to r)) \to (r \to p)$.

Solution. *(a)* Suppose the formula is F. Then r is F, and the hypothesis is T, so each of $p \vee q$, $p \to r$, $q \to r$ is T (V3). With r false, $p \to r$ true forces $p = \mathrm{F}$, and $q \to r$ true forces $q = \mathrm{F}$. Then $p \vee q$ is F, a contradiction. So the formula is false in no row. It is the table's version of proof by cases: if r follows from each of p and q, and one of them holds, then r holds.

(b) Suppose it is F. Then $p \to (q \to r)$ is T and $(p \to q) \to (p \to r)$ is F; the latter forces $p \to q$ to be T and $p \to r$ to be F, and the last forces $p = \mathrm{T}$, $r = \mathrm{F}$. Then $p \to q$ true with p true forces $q = \mathrm{T}$. Now $p \to (q \to r)$ has a true hypothesis and the conclusion $q \to r$, which is F (true q, false r), so it is F, a contradiction. False in no row.

(c) Suppose it is F. Then $(p \vee q) \to r$ is F: $r = \mathrm{F}$ and $p \vee q$ is T. And $(p \to r) \vee (q \to r)$ is T: with r false, $p \to r$ is T exactly when $p = \mathrm{F}$, and $q \to r$ exactly when $q = \mathrm{F}$; so at least one of p, q is F. Together with $p \vee q$ true: exactly one of p, q is T. This leaves a choice, so split. The formula is false in exactly two rows: $p = \mathrm{T}, q = \mathrm{F}, r = \mathrm{F}$ and $p = \mathrm{F}, q = \mathrm{T}, r = \mathrm{F}$. Check the first: $q \to r$ is T, so the hypothesis is T; $p \vee q$ is T and r is F, so the conclusion is F.

(d) Suppose it is F. Then $r \to p$ is F: $r = \mathrm{T}$, $p = \mathrm{F}$. The hypothesis must be T: $p \to q$ is T because p is F, and $q \to r$ is T because r is T, whatever q is. Nothing forces q. So the formula is false in exactly two rows: $p = \mathrm{F}$, $r = \mathrm{T}$, and q either value.

Exercise. B2: Counting atoms, the last step blank

Call a formula **binary-built** if it is built from atoms by the four binary connectives only, so that neither $\neg$ nor $\bot$ occurs in it; the immediate parts of a binary-built formula are again binary-built. Define by recursion (Theorem 4.3.4) the number $\mathrm{occ}(\varphi)$ of atom occurrences and the number $\mathrm{bin}(\varphi)$ of binary connectives:

$$\begin{aligned} \mathrm{occ}(p) &= 1, \quad \mathrm{occ}(\varphi \circ \psi) = \mathrm{occ}(\varphi) + \mathrm{occ}(\psi), \\ \mathrm{bin}(p) &= 0, \quad \mathrm{bin}(\varphi \circ \psi) = \mathrm{bin}(\varphi) + \mathrm{bin}(\psi) + 1, \end{aligned}$$

for every atom p and every binary connective $\circ$. (For $\bot$ and $\neg$ give both functions any values;

they play no part below.) Here is a proof, by Theorem 4.3.1, that every binary-built formula φ has

$$\operatorname{occ}(\varphi) = \operatorname{bin}(\varphi) + 1.$$

The property is $C(\varphi)$: *if φ is binary-built, then* $\operatorname{occ}(\varphi) = \operatorname{bin}(\varphi) + 1$.

$\bot$ and negation. Neither $\bot$ nor a formula $\neg\varphi$ is binary-built, so C holds for them: its "if" is never met.

Atoms. $\operatorname{occ}(p) = 1 = 0 + 1 = \operatorname{bin}(p) + 1$.

Conjunction. Assume $C(\varphi)$ and $C(\psi)$, and let $\varphi \wedge \psi$ be binary-built. Then φ and ψ are binary-built, so

$$\begin{aligned}\operatorname{occ}(\varphi \wedge \psi) &= \operatorname{occ}(\varphi) + \operatorname{occ}(\psi) = \big(\operatorname{bin}(\varphi) + 1\big) + \big(\operatorname{bin}(\psi) + 1\big) \\ &= \big(\operatorname{bin}(\varphi) + \operatorname{bin}(\psi) + 1\big) + 1 = \operatorname{bin}(\varphi \wedge \psi) + 1.\end{aligned}$$

Disjunction and biconditional. The computation is the same with $\vee$, and with $\leftrightarrow$, in place of $\wedge$.

Implication. _____

Name the move, then complete it.

Solution. The move is **induction on the shape of a formula** (Theorem 4.3.1): the case for a compound formula uses the property for its immediate parts, and nothing else.

Implication. Assume $C(\varphi)$ and $C(\psi)$, and let $\varphi \to \psi$ be binary-built. Its immediate parts φ and ψ are binary-built, so $\operatorname{occ}(\varphi) = \operatorname{bin}(\varphi) + 1$ and $\operatorname{occ}(\psi) = \operatorname{bin}(\psi) + 1$. Hence

$$\begin{aligned}\operatorname{occ}(\varphi \to \psi) &= \operatorname{occ}(\varphi) + \operatorname{occ}(\psi) = \big(\operatorname{bin}(\varphi) + 1\big) + \big(\operatorname{bin}(\psi) + 1\big) \\ &= \big(\operatorname{bin}(\varphi) + \operatorname{bin}(\psi) + 1\big) + 1 = \operatorname{bin}(\varphi \to \psi) + 1,\end{aligned}$$

which is $C(\varphi \to \psi)$. All clauses of Definition 4.1.1 are covered, so Theorem 4.3.1 gives $C(\varphi)$ for every formula, and in particular the equation for every binary-built formula.

For instance $(p \to q) \wedge (p \vee r)$ has four atom occurrences (p counted twice) and three binary connectives.

Exercise. B3: A table for a named pattern

Make the truth table of $((p \vee q) \wedge \neg p) \to q$, and say which derived rule of Chapter 2 it corresponds to, as Example (The truth table of modus ponens) did for $\to$E.

Solution.

p	q	$p \vee q$	$\neg p$	$(p \vee q) \wedge \neg p$	$((p \vee q) \wedge \neg p) \to q$
T	T	T	F	F	T
T	F	T	F	F	T
F	T	T	T	T	T
F	F	F	T	F	T

True in every row. The only row in which both $p \vee q$ and $\neg p$ are true is row 3, and there q is true. This is disjunctive syllogism, DS (Proposition 2.3.4): from $P \vee Q$ and $\neg P$, infer Q.

Exercise. B4: Exactly one is true

An integer n is such that **exactly one** of these three statements is true.

- (A) n is even.
- (B) If n is even, then n is divisible by 4.
- (C) n is divisible by 4, or n is odd.

With the atom p for "n is even" and the atom q for "n is divisible by 4", write (A), (B), (C) as formulas, find every row on p, q in which exactly one of them is true, and say what n can be.

Solution.

(A) is p, (B) is $p \to q$, and (C) is $q \vee \neg p$ ("n is odd" is "n is not even").

p	q	(A)	(B)	(C)	true
T	T	T	T	T	three
T	F	T	F	F	one
F	T	F	T	T	two
F	F	F	T	T	two

In the row TF, (B) has a true hypothesis and a false conclusion, and (C) has two false disjuncts. In the rows with p false, (B) has a false hypothesis and (C) the true disjunct $\neg p$. So only the row $p = \text{T}$, $q = \text{F}$ makes exactly one statement true: n is even and not divisible by 4, for instance 2, 6 or -10.

The row FT is impossible for an integer, since a number divisible by 4 is even; it does not matter, as it fails the puzzle anyway. (B) and (C) agree in every row, and Section 4.5 explains why.

Marking scheme. (10 points)

- 3: the three formulas.
- 3: the four rows evaluated.
- 2: the conclusion about n.
- 2: an example.

Exercise. B5: Which primes divide n?

Of an integer n we know that all four of these claims hold.

1. If 2 divides n, then 3 does not.
2. 3 divides n, or 5 divides n.
3. If 5 divides n, then 2 does.

4. 2 and 5 do not both divide n.

With the atoms p, q, r for "$2 \mid n$", "$3 \mid n$", "$5 \mid n$", find every row making all four claims true by working backwards, and give two such integers n. Why may the three atoms be treated as independent here?

Solution. The claims are $p \to \neg q$, $q \vee r$, $r \to p$ and $\neg(p \wedge r)$.

If r were T, claim 3 would force $p = \mathrm{T}$ (a true implication with a true hypothesis has a true conclusion), and then $p \wedge r$ would be T, against claim 4. So $r = \mathrm{F}$. Then claim 2 forces $q = \mathrm{T}$, and claim 1 forces $p = \mathrm{F}$: with p true, its conclusion $\neg q$ would be false. Check: the row $p = \mathrm{F}$, $q = \mathrm{T}$, $r = \mathrm{F}$ makes all four claims true, and since every step was forced, it is the only such row.

So 3 divides n, and 2 and 5 do not: $n = 3$, 9 or 21, for instance.

Independence: each of the eight rows is the row of some integer. The integers $30, 6, 10, 2, 15, 3, 5, 1$ give the rows TTT, TTF, TFT, TFF, FTT, FTF, FFT, FFF, in that order. So arithmetic rules out no row in advance, and the table's answer is the answer about integers.

Marking scheme. (10 points)

- 2: the four formulas.
- 4: the forced steps, each with its clause.
- 2: the check and two values of n.
- 2: independence.

4.4.7.3 C. Going deeper

Exercise. C1: A formula built from biconditionals

Call a formula **$\leftrightarrow$-built** if it is built from atoms by $\leftrightarrow$ alone. For a valuation v and a formula φ, let $f_v(\varphi)$ be the number of occurrences of atoms in φ that are false under v, counted with repetition.

(a) Prove by structural induction that every $\leftrightarrow$-built φ is true under v exactly when $f_v(\varphi)$ is even.

(b) Deduce that $(p \leftrightarrow q) \leftrightarrow p$ has the same value as q under every valuation.

Solution (Hint 1). Put "if φ is $\leftrightarrow$-built" inside the property, as in Exercise B2.

Solution (Hint 2). In the step, $f_v(\varphi \leftrightarrow \psi) = f_v(\varphi) + f_v(\psi)$. When is a sum of two numbers even?

Solution. (a) The property is $C(\varphi)$: *if φ is $\leftrightarrow$-built, then for every valuation v: $v(\varphi) = \mathrm{T}$ iff $f_v(\varphi)$ is even.* The cases $\bot$, $\neg$ and the other three binary connectives are vacuous, since such formulas are not $\leftrightarrow$-built.

Atoms. If $v(p) = \mathrm{T}$, then $f_v(p) = 0$, which is even; if $v(p) = \mathrm{F}$, then $f_v(p) = 1$, which is odd. Either way, $v(p) = \mathrm{T}$ iff $f_v(p)$ is even.

Step. Assume $C(\varphi)$ and $C(\psi)$, and let $\varphi \leftrightarrow \psi$ be $\leftrightarrow$-built. Its parts are $\leftrightarrow$-built, so the hypotheses apply to them. By (V6), $v(\varphi \leftrightarrow \psi) = \mathrm{T}$ iff $v(\varphi) = v(\psi)$, iff $f_v(\varphi)$ and $f_v(\psi)$ are both even or both odd (by the hypotheses), iff their sum $f_v(\varphi \leftrightarrow \psi)$ is even.

Every clause of Definition 4.1.1 is covered, so Theorem 4.3.1 gives $C(\varphi)$ for every formula.

(b) In $(p \leftrightarrow q) \leftrightarrow p$, the atom p occurs twice and q once. So f_v counts 2 if p is false (0 otherwise), plus 1 if q is false (0 otherwise), which is even exactly when q is true. By (a), the formula is true exactly when q is.

Marking scheme. (10 points)

- 2: the property stated for every formula, with the restriction inside and "for every v".
- 1: the vacuous cases named.
- 2: the atom case, both values.
- 4: the step, with (V6) and the parity of a sum.
- 1: (b).

Fatal errors (0 for the step): a step that uses the hypothesis for one part only; fixing v outside the property and then using another valuation.

Exercise. C2: Find the error, used later

Here is a claim and a "proof".

> **Claim.** Let v_{T} be the all-T valuation. Every formula built from atoms by $\wedge$, $\vee$ and $\rightarrow$ only is true under v_{T}.
>
> **Proof.** By Theorem 4.3.1, with the property "if φ is built from atoms by $\wedge$, $\vee$, $\rightarrow$ only, then $v_{\mathrm{T}}(\varphi) = \mathrm{T}$". The formulas $\bot$, $\neg\varphi$ and $\varphi \leftrightarrow \psi$ are not built in this way, so the property holds for them. If a formula built in this way is $\varphi \wedge \psi$, $\varphi \vee \psi$ or $\varphi \rightarrow \psi$, then φ and ψ are built in this way too, so the hypothesis applies to them. (1) For an atom p, $v_{\mathrm{T}}(p) = \mathrm{T}$ by the definition of v_{T}. (2) If $v_{\mathrm{T}}(\varphi) = \mathrm{T}$ and $v_{\mathrm{T}}(\psi) = \mathrm{T}$, then $v_{\mathrm{T}}(\varphi \wedge \psi) = \mathrm{T}$ by (V3). (3) If $v_{\mathrm{T}}(\varphi) = \mathrm{T}$ and $v_{\mathrm{T}}(\psi) = \mathrm{T}$, then $v_{\mathrm{T}}(\varphi \vee \psi) = \mathrm{T}$ by (V4), since at least one part is true. (4) If $v_{\mathrm{T}}(\varphi) = \mathrm{T}$ and $v_{\mathrm{T}}(\psi) = \mathrm{T}$, then $v_{\mathrm{T}}(\varphi \rightarrow \psi) = \mathrm{T}$ by (V5), because $v_{\mathrm{T}}(\varphi) = \mathrm{F}$ and an implication with a false hypothesis is true. This covers every clause, and proves the claim.

(a) Is the claim true?

(b) Name the line that fails and say exactly what is wrong with it.

(c) Give a formula on which the reason given in that line is false.

(d) Repair the line.

Solution. (a) Yes. The claim is true, and the repaired proof in (d) proves it. So no formula is a counterexample to the claim; the error is in the argument only.

(b) Line (4). Its hypothesis states $v_{\mathrm{T}}(\varphi) = \mathrm{T}$, and its reason states $v_{\mathrm{T}}(\varphi) = \mathrm{F}$. The reason contradicts the induction hypothesis it is supposed to use: it quotes the vacuous row of (V5), which does not apply here. The conclusion $v_{\mathrm{T}}(\varphi \rightarrow \psi) = \mathrm{T}$ is correct, but not for the reason

given.

(c) $p \to q$. Under v_{T}, $v_{\mathrm{T}}(p) = \mathrm{T}$, so the stated reason "$v_{\mathrm{T}}(p) = \mathrm{F}$" is false, although $v_{\mathrm{T}}(p \to q) = \mathrm{T}$ holds.

(d) (4) If $v_{\mathrm{T}}(\varphi) = \mathrm{T}$ *and* $v_{\mathrm{T}}(\psi) = \mathrm{T}$*, then* $v_{\mathrm{T}}(\varphi \to \psi) = \mathrm{T}$ *by (V5), because the conclusion* ψ *is true, so we are not in the one row (true hypothesis, false conclusion) where an implication is false.* With this line the induction is correct, and the claim follows. Section 4.9 uses this fact to show that $\wedge$, $\vee$ and $\to$ cannot express negation.

Marking scheme. (10 points)

- 2: the claim is true (with the reason: the repaired proof).
- 3: names line (4), and says its reason contradicts the hypothesis $v_{\mathrm{T}}(\varphi) = \mathrm{T}$ of that very case.
- 2: a formula, such as $p \to q$, on which the stated reason fails.
- 3: the repair, citing the row of (V5) that actually applies (conclusion true).

Fatal errors (0 for the step): calling the claim false; locating the error in line (3) (at least one part true is correct) or in the handling of $\bot$, $\neg$, $\leftrightarrow$.

Exercise. C3: The game and the clauses

Fix a valuation v. For formulas built from atoms by $\neg$, $\wedge$, $\vee$ alone, define by recursion (Theorem 4.3.4, applied to the pair of truth values of the two statements as one function, since each clause for $\neg$ calls on the other statement; give $\bot$, $\to$ and $\leftrightarrow$ any values, they play no part) two statements, $\mathrm{Ver}(\varphi)$ ("the verifier can always win at φ") and $\mathrm{Fal}(\varphi)$ ("the falsifier can always win at φ"):

φ	$\mathrm{Ver}(\varphi)$ holds iff	$\mathrm{Fal}(\varphi)$ holds iff
an atom p	$v(p) = \mathrm{T}$	$v(p) = \mathrm{F}$
$\neg\theta$	$\mathrm{Fal}(\theta)$	$\mathrm{Ver}(\theta)$
$\theta \wedge \chi$	$\mathrm{Ver}(\theta)$ and $\mathrm{Ver}(\chi)$	$\mathrm{Fal}(\theta)$ or $\mathrm{Fal}(\chi)$
$\theta \vee \chi$	$\mathrm{Ver}(\theta)$ or $\mathrm{Ver}(\chi)$	$\mathrm{Fal}(\theta)$ and $\mathrm{Fal}(\chi)$

(a) Say in words why each clause matches the rules of the game.

(b) Prove by structural induction that for every such formula φ, $\mathrm{Ver}(\varphi)$ holds iff $v(\varphi) = \mathrm{T}$, and $\mathrm{Fal}(\varphi)$ holds iff $v(\varphi) = \mathrm{F}$.

(c) Why must the two statements of (b) be proved together?

Solution (Hint 1). Put "if φ is built from $\neg$, $\wedge$, $\vee$" inside the property.

Solution (Hint 2). The step for $\neg$ needs the claim about Fal for the part.

Solution. (a) At an atom the game ends, and the value of the atom decides the winner. At $\wedge$ the falsifier chooses, so the verifier wins only if she can win at both parts, whichever he picks, and the falsifier wins if he can win at one of them, by picking it. At $\vee$ the verifier chooses, so the roles are reversed: she needs one part, he needs both. At $\neg$ the players swap, so winning as verifier at $\neg\theta$ is winning as falsifier at θ, and the other way round.

(b) The property is $C(\varphi)$: *if φ is built from $\neg$, $\wedge$, $\vee$, then (*$\mathrm{Ver}(\varphi)$ *iff $v(\varphi) = \mathrm{T}$) and (*$\mathrm{Fal}(\varphi)$ *iff $v(\varphi) = \mathrm{F}$).* For $\bot$, $\to$ and $\leftrightarrow$ it is vacuous. For an atom both halves are the definition.

Negation. Assume $C(\theta)$. Then $\mathrm{Ver}(\neg\theta)$ iff $\mathrm{Fal}(\theta)$ iff $v(\theta) = \mathrm{F}$ iff $v(\neg\theta) = \mathrm{T}$, by (V2). Likewise $\mathrm{Fal}(\neg\theta)$ iff $\mathrm{Ver}(\theta)$ iff $v(\theta) = \mathrm{T}$ iff $v(\neg\theta) = \mathrm{F}$.

Conjunction. Assume $C(\theta)$ and $C(\chi)$. $\mathrm{Ver}(\theta \wedge \chi)$ iff both $\mathrm{Ver}(\theta)$ and $\mathrm{Ver}(\chi)$, iff both parts are true, iff $v(\theta \wedge \chi) = \mathrm{T}$, by (V3). $\mathrm{Fal}(\theta \wedge \chi)$ iff Fal holds of one part, iff one part is false, iff $v(\theta \wedge \chi) = \mathrm{F}$.

Disjunction. The same with (V4): Ver iff one part is true, Fal iff both parts are false.

Every clause of Definition 4.1.1 is covered, so Theorem 4.3.1 gives $C(\varphi)$ for every formula.

(c) The $\neg$ step for Ver uses the hypothesis about Fal for the part, and the $\neg$ step for Fal uses the one about Ver. Either half alone does not pass through the step, so the property must carry both: Section 4.2's move, proving more in order to prove less.

Marking scheme. (10 points)

- 2: (a).
- 2: the property, with the restriction and both halves.
- 1: the vacuous cases and the atoms.
- 2: the $\neg$ step.
- 2: the $\wedge$ and $\vee$ steps.
- 1: (c).

Fatal error (0 for the step): proving only the Ver half and using the Fal half in the $\neg$ step without proof.

4.5 Tautologies, Entailment and Equivalence

Section 4.4 gave every formula a value under every valuation, and collected the values in a truth table. One formula at a time, though, is not what Chapters 1 and 2 were about. They were about **sequents**: premises on the left, a conclusion on the right, and the claim that the conclusion can be derived. This section defines the counterpart of a sequent in terms of truth, $\Gamma \models \psi$: every valuation that makes the premises true makes the conclusion true. On the way it names the formulas that are true in every row, and the pairs of formulas that are true in the same rows. It then proves two lemmas about changing a formula, by replacing one part or by putting formulas in for its atoms, that later sections use, and ends by sorting the section's questions by whether one row or every row settles them, and by setting $\models$ beside $\vdash$.

4.5.1 Tautologies and contradictions

Some tables of Section 4.4 had a last column made entirely of T: the formula $((p \to q) \wedge p) \to q$ of Example (The truth table of modus ponens), and the biconditional of Example (An implication and its contrapositive). Their truth does not depend on the situation. One table in Exercise A1 there, that of $p \leftrightarrow \neg p$, was all F. These extremes deserve names.

A tautology is true for reasons of form alone: whatever the atoms mean, it comes out true.

Definition 4.5.1. Tautology and Contradiction

A formula φ is a **tautology** if $v(\varphi) = \mathrm{T}$ for **every** valuation v. It is a **contradiction** if $v(\varphi) = \mathrm{F}$ for **every** valuation v.

By Lemma 4.4.1, the value of φ under a valuation depends only on the row the valuation gives to the atoms of φ. So φ is a tautology exactly when its truth table has T in every row, and a contradiction exactly when it has F in every row. Infinitely many valuations, finitely many rows to check.

Examples, simplest first.

- $p \to p$ and $p \vee \neg p$ are tautologies: in the row $p = \mathrm{T}$ both are true, and in the row $p = \mathrm{F}$, $p \to p$ has a false hypothesis and $\neg p$ is true.
- $p \wedge \neg p$ and $p \leftrightarrow \neg p$ are contradictions: in each row one of p, $\neg p$ is false, and the two sides of the biconditional differ.
- *Nonexample.* $p \to q$ is neither: it is T in the row TT and F in the row TF. Most formulas are like this. Changing one symbol of the tautology $p \to p$, to get $p \to q$, loses the property: the rows in which the hypothesis and conclusion differ are no longer excluded by their being the same atom. A formula that is true in some rows and false in others, such as $p \to q$, is called **contingent**: every formula is a tautology, a contradiction, or contingent.

↻ **Second reading** · *Formulas with no atoms*

A formula with no atoms, such as $\bot$ or $\neg\bot$, has the same value under every valuation (its table has one row). So it is either a contradiction, like $\bot$, or a tautology, like $\neg\bot$. This degenerate case is the only one in which "not a tautology" forces "a contradiction".

Lists of formulas need a word too, because a list of premises can fail to be true together even when each premise alone can be true.

Definition 4.5.2. Satisfiable

A formula φ is **satisfiable** if **there is** a valuation v with $v(\varphi) = \mathrm{T}$. A list of formulas $\Gamma = \varphi_1, \ldots, \varphi_n$, where $n \geq 0$, is **satisfiable** if **there is** a valuation that makes **every** φ_i true. A formula or list that is not satisfiable is **unsatisfiable**.

In words: a formula is satisfiable if some row makes it true, and a list is satisfiable if some **single** row makes all its members true at once. So a formula is unsatisfiable exactly when it is a contradiction. The **empty list** ($n = 0$) is satisfiable: every valuation makes all of its members

true, since it has none to check.

The list $p, \neg p$ is unsatisfiable, although each of its members is satisfiable: a row making p true makes $\neg p$ false. Satisfiability of a list is about a common row, not about each formula separately. And the list $p \to q,\ q \to r,\ r \to \neg p$ is satisfiable, by the row with p false and q, r true (the first two have true conclusions, and the third has conclusion $\neg p$, which is true); in fact p is false in every row that makes all three true.

Tautologies and unsatisfiable formulas are two sides of one coin. For every formula ψ,

$$\psi \text{ is a tautology} \iff \neg\psi \text{ is unsatisfiable.}$$

Indeed, by (V2) of Definition 4.4.2, $v(\neg\psi) = \mathrm{T}$ exactly when $v(\psi) = \mathrm{F}$. So "no valuation makes $\neg\psi$ true" says "no valuation makes ψ false", which is "every valuation makes ψ true". Sections 4.6 and 4.9 use it to pass between the two kinds of question.

Remark. **Satisfiable and consistent.** Section 3.2 called a list Γ *inconsistent* if $\Gamma \vdash \bot$ (Definition 3.2.1), and *consistent* otherwise: a word about derivations. *Satisfiable* is its twin about truth: some row makes every member of the list true. Nothing yet connects the two words. Section 4.7 proves that a satisfiable list is consistent, and Section 4.8 the converse; until then we keep them apart. The counterexample set lives in both worlds: $\Gamma \vdash \psi$ exactly when $\Gamma, \neg\psi$ is inconsistent (Proposition 3.2.2 (b)), and $\Gamma \models \psi$ exactly when $\Gamma, \neg\psi$ is unsatisfiable (below).

Warning. "Not a tautology" does not mean "a contradiction". The formula $p \to q$ is neither: some rows make it true and some false. To show a formula is **not** a tautology, one row making it false is enough; to show it **is** a contradiction, every row must make it false.

4.5.2 Semantic entailment

Now the counterpart of a sequent. A derivation from premises Γ to a conclusion ψ is a promise that ψ follows from Γ. In terms of truth, "follows" should mean: no situation makes the premises true and the conclusion false. We want to prove, in Sections 4.7 and 4.8, that the rules of Chapters 1 and 2 derive exactly the sequents that pass this test; for that we need a precise name for the test.

An argument is valid when its conclusion is true in every row in which all its premises are true.

Definition 4.5.3. Semantic Entailment

Let $\Gamma = \varphi_1, \ldots, \varphi_n$ be a list of formulas, where $n \geq 0$, and let ψ be a formula. We say Γ **semantically entails** ψ, and write

$$\Gamma \models \psi,$$

if **every** valuation that makes **all** of $\varphi_1, \ldots, \varphi_n$ true also makes ψ true. We then also call the argument from Γ to ψ **valid**. We write $\varphi_1, \ldots, \varphi_n \models \psi$ for $\Gamma \models \psi$, and $\models \psi$ when Γ is empty. If $\Gamma \models \psi$ fails, we write $\Gamma \not\models \psi$, and a valuation that makes every formula of Γ true and ψ false is a **refuting valuation**, or **refuting row**.

Clause by clause. "Every valuation": the condition is checked in all situations, not in the actual

one; a valid argument cannot be spoiled by what the atoms happen to mean. "That makes all of $\varphi_1, \ldots, \varphi_n$ true": rows in which some premise is false are ignored, whatever the conclusion does there. "Also makes ψ true": in the remaining rows, the conclusion must hold. In symbols, for every valuation v,

$$\big(v(\varphi_1) = \mathrm{T} \text{ and } \ldots \text{ and } v(\varphi_n) = \mathrm{T}\big) \;\Rightarrow\; v(\psi) = \mathrm{T}.$$

And $\Gamma \not\models \psi$ says: **there is** a valuation making every premise true and the conclusion false. One refuting row is enough.

Checking it. By Lemma 4.4.1, whether a valuation makes the premises and the conclusion true depends only on its row on the atoms occurring in Γ and ψ. So $\Gamma \models \psi$ can be decided by one truth table on those atoms (Definition 4.4.3 allows extra atoms in the list): find the rows in which every premise is T, and look at the conclusion there.

First examples. Row 1 of Example (The truth table of modus ponens) is the only row making $p \to q$ and p true, and q is true there, so $p \to q,\ p \models q$. By Example (An implication and its contrapositive), $p \to q \models \neg q \to \neg p$ and $\neg q \to \neg p \models p \to q$, since the two columns are equal.

The empty list. When Γ is empty, every valuation makes "all of Γ" true, so $\models \psi$ says that every valuation makes ψ true: $\models \psi$ holds exactly when ψ is a tautology.

Unsatisfiable premises. If Γ is unsatisfiable, no valuation makes all of Γ true, so the condition has no row to check and holds: $\Gamma \models \psi$ for **every** formula ψ. For example $p, \neg p \models q$ and $\bot \models q$. This is the table's version of Proposition 2.1.1, $P, \neg P \vdash Q$, and of $\bot$E.

More premises. If $\Gamma \models \psi$, and Δ is a list containing every formula of Γ (and perhaps others), then $\Delta \models \psi$: a valuation making all of Δ true makes all of Γ true, and so makes ψ true. Adding premises never destroys an entailment. In the same way, the order of the premises and repetitions among them do not matter: only which formulas occur in the list does.

↻ **Second reading** · *The bookkeeping of entailment*

Four facts about $\models$ need nothing but the definition. *Reflexivity:* $\Gamma \models \varphi$ whenever φ is a member of Γ, since a row making all of Γ true makes φ true. *Weakening:* adding premises keeps an entailment, as just shown. *Order and repetition:* rearranging Γ, or repeating a member, changes nothing. *Cut:* if $\Gamma \models \varphi$ and $\Gamma, \varphi \models \psi$, then $\Gamma \models \psi$. Indeed, let a row make every member of Γ true. By the first entailment it makes φ true, so it makes every member of Γ, φ true, and by the second it makes ψ true. Section 2.3 proved weakening and cut for $\vdash$ (Lemma 2.3.3) by pasting derivations, and its weakening, which asks only that every member of Γ be a member of Δ, covers order and repetition; reflexivity for $\vdash$ is the derivation whose only line is the premise φ. Here each fact is one line about rows. Section 4.8 sets the two lists side by side.

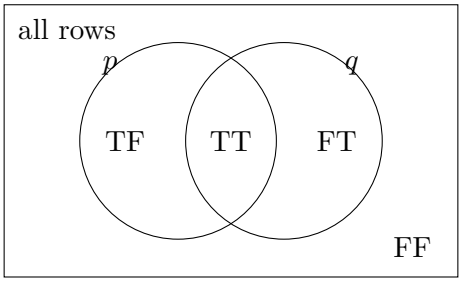
A test by satisfiability. Finally, for every list Γ and formula ψ,

$$\Gamma \models \psi \iff \text{the list } \Gamma, \neg\psi \text{ is unsatisfiable.}$$

Both sides say that no valuation makes every formula of Γ true and ψ false, because $v(\neg\psi) = \mathrm{T}$ exactly when $v(\psi) = \mathrm{F}$. With Γ empty this is the equivalence between tautologies and unsatisfiable negations above.

The list $\Gamma, \neg\psi$ is the **counterexample set** of the claim (Definition 1.1.3): an argument is valid exactly when its counterexample set cannot be made true by any row.

Rows as regions. A table on p and q has four rows, and they can be drawn as the four regions of a picture.



Each row of the table on p, q is a region. A formula marks the rows where it is true: $p \wedge q$ the lens, $p \vee q$ both circles, $\neg p$ everything outside p's circle. $\Gamma \models \psi$ says that the region where all of Γ is true lies inside the region of ψ: $p \wedge q \models p \vee q$.

Equivalent formulas mark the same region, a tautology marks the whole rectangle, and an unsatisfiable list marks nothing in common.

Quick check. Does $p \wedge q \models p \vee r$ hold? Does $p \vee r \models p \wedge q$?

Solution. The first holds: a row making $p \wedge q$ true makes p true, so $p \vee r$ is true. The second fails, for instance at $p = \text{F}$, $r = \text{T}$ (any q), which makes $p \vee r$ true and $p \wedge q$ false.

Example. Affirming the consequent is not valid

Show that $p \to q,\ q \not\models p$.

Solution. We need one row making both premises true and the conclusion false. The conclusion p must be F, and the premise q must be T. In the row $p = \text{F}$, $q = \text{T}$, the premise $p \to q$ is T, since its hypothesis is false. So this row makes $p \to q$ and q true and p false: it is a refuting row.

The pattern is the one of Section 1.3, "from $\varphi \to \psi$ and ψ, infer φ", which is not an instance of $\to$E. Section 1.1's Quick check gave the row a face: with P: "n is divisible by 4" and Q: "n is even", at $n = 2$ the premises $P \to Q$ and Q are true and P is false. It is also the step that would repair Argument B of Section 1.1: from "if n is even, then n^2 is even" and "n^2 is even", conclude "n is even". The conclusion of that instance happens to be true for every integer n, as Argument A proves; the refuting row shows that the **form** does not guarantee it, so the step proves nothing. Section 4.7 upgrades this row from "not valid" to "not derivable by any derivation".

A valid argument with several premises is checked in the same way, and working backwards from F often settles it without a table.

Example. Two cases, two conclusions

Show that $p \vee q,\ p \to r,\ q \to s \models r \vee s$.

Solution. Suppose a valuation v makes all three premises true and the conclusion false; we derive a contradiction. Since $r \vee s$ is F, both r and s are F (V4). Since $p \to r$ is T and r is F, the hypothesis p is F (V5). In the same way $q \to s$ forces q to be F. But then $p \vee q$ is F, contradicting the first premise. So no refuting valuation exists, and the entailment holds.

The search replaced a table of $2^4 = 16$ rows by four forced steps. The entailment is the table's picture of a proof by cases with $\vee$E: in the case p, the second premise gives r, and $\vee$I gives $r \vee s$; in the case q, the third gives s, and $\vee$I again.

Changing the grouping of a formula can change what it entails.

Example. One grouping entails the other, not conversely

Decide whether $(p \to q) \to r \models p \to (q \to r)$, and whether $p \to (q \to r) \models (p \to q) \to r$.

Solution. The first holds. Suppose a valuation makes $(p \to q) \to r$ true and $p \to (q \to r)$ false. The latter forces $p =$ T and $q \to r$ false, so $q =$ T and $r =$ F. Then $p \to q$ is T, and $(p \to q) \to r$ has a true hypothesis and a false conclusion, so it is F, a contradiction.

The second fails. Take $p =$ F and $r =$ F, with q either value. Then $p \to (q \to r)$ is T, since its hypothesis is false. But $p \to q$ is T for the same reason, so $(p \to q) \to r$ is T $\to$ F, which is F. These two rows are refuting rows.

So the two groupings are not interchangeable, which is why Chapter 1 fixed a convention ($\to$ groups to the right) and why the convention matters.

Warning. $\models$ is not $\to$. The formula $\varphi \to \psi$ is a formula: it has a value in each row and can be a part of larger formulas. The statement $\varphi \models \psi$ is not a formula at all. It is a claim **about** two formulas, true or false once and for all, like $\varphi \vdash \psi$. They are related (Exercise B3 below asks you to prove that $\varphi \models \psi$ holds exactly when $\varphi \to \psi$ is a tautology), but writing $p \models q \models r$, or $\neg(p \models q)$ inside a formula, mixes the two levels and means nothing.

Quick check. Is $p, \neg p \models q$? Is $p \to q,\ \neg p \models \neg q$?

Solution. The first holds: the list $p, \neg p$ is unsatisfiable, so there is no row to check. The second fails: the row $p =$ F, $q =$ T makes $p \to q$ true (false hypothesis) and $\neg p$ true, and $\neg q$ false. It is the pattern "denying the hypothesis", the mirror image of affirming the consequent.

4.5.3 Semantic equivalence

Two formulas can be different strings and still say the same thing in every situation, as $p \to q$ and $\neg q \to \neg p$ did in Example (An implication and its contrapositive). Chapter 2 had a word for formulas that derive each other, provable equivalence (Definition 2.4.2). Here is the counterpart in terms of truth.

Two formulas are equivalent when they have the same truth table.

Definition 4.5.4. Semantic Equivalence

Formulas φ and ψ are **semantically equivalent**, written $\varphi \equiv \psi$, if $\varphi \models \psi$ **and** $\psi \models \varphi$.

In words, each entails the other. Equivalently, $v(\varphi) = v(\psi)$ for **every** valuation v. Indeed, if $\varphi \equiv \psi$ and $v(\varphi) = \mathrm{T}$, then $v(\psi) = \mathrm{T}$ by $\varphi \models \psi$; if $v(\varphi) = \mathrm{F}$, then $v(\psi)$ cannot be T, since $\psi \models \varphi$ would make $v(\varphi) = \mathrm{T}$. The converse is plain. By (V6) of Definition 4.4.2, "$v(\varphi) = v(\psi)$ for every v" says that $\varphi \leftrightarrow \psi$ is true under every v, so

$$\varphi \equiv \psi \iff \models \varphi \leftrightarrow \psi.$$

From "same value under every valuation" it follows at once that $\equiv$ is reflexive, symmetric, and transitive: $\varphi \equiv \varphi$; if $\varphi \equiv \psi$, then $\psi \equiv \varphi$; if $\varphi \equiv \psi$ and $\psi \equiv \chi$, then $\varphi \equiv \chi$.

Examples: $p \wedge q \equiv q \wedge p$ (the clause (V3) does not care about the order); $p \equiv \neg\neg p$ (two flips); $p \vee \neg p \equiv \neg\bot$ (both are true under every valuation). The degenerate case: any two tautologies are equivalent, and so are any two contradictions. *Nonexample:* $p \wedge q$ and p are not equivalent. One direction holds, $p \wedge q \models p$, but the other fails in the row $p = \mathrm{T}$, $q = \mathrm{F}$. Entailment in one direction is not equivalence. Nor is equivalence identity: $p \wedge q$ and $q \wedge p$ are different formulas, with different parse trees, and $\to$E treats them as different (Section 1.3).

The equivalences used most often are collected below. They are stated for arbitrary formulas φ, ψ, χ, not just atoms, since Section 4.9 applies them to compound formulas.

Proposition 4.5.1. Standard Equivalences

For all formulas φ, ψ and χ:

(a) (contrapositive) $\varphi \to \psi \equiv \neg\psi \to \neg\varphi$;
(b) (De Morgan) $\neg(\varphi \wedge \psi) \equiv \neg\varphi \vee \neg\psi$ and $\neg(\varphi \vee \psi) \equiv \neg\varphi \wedge \neg\psi$;
(c) (implication as a disjunction) $\varphi \to \psi \equiv \neg\varphi \vee \psi$;
(d) (double negation) $\neg\neg\varphi \equiv \varphi$;
(e) (biconditional) $\varphi \leftrightarrow \psi \equiv (\varphi \to \psi) \wedge (\psi \to \varphi)$;
(f) (commutativity) $\varphi \wedge \psi \equiv \psi \wedge \varphi$ and $\varphi \vee \psi \equiv \psi \vee \varphi$;
(g) (distributivity) $\varphi \wedge (\psi \vee \chi) \equiv (\varphi \wedge \psi) \vee (\varphi \wedge \chi)$ and $\varphi \vee (\psi \wedge \chi) \equiv (\varphi \vee \psi) \wedge (\varphi \vee \chi)$;
(h) (associativity) $(\varphi \wedge \psi) \wedge \chi \equiv \varphi \wedge (\psi \wedge \chi)$ and $(\varphi \vee \psi) \vee \chi \equiv \varphi \vee (\psi \vee \chi)$.

Idea. For atoms these are truth-table checks. For arbitrary formulas the same tables work, with φ, ψ, χ in the place of atoms: under a fixed valuation, each of φ, ψ, χ has some value, and both sides are computed from those values by the clauses (V2)–(V6), exactly as in a table whose first columns are φ, ψ, χ.

Proof. Fix a valuation v, and let $a = v(\varphi)$, $b = v(\psi)$, $c = v(\chi)$. Each side of each equivalence is built from φ, ψ, χ by connectives, so by (V2)–(V6) of Definition 4.4.2 its value under v is computed from a, b, c by the clauses of those connectives. The combination (a, b), or (a, b, c), is one row of the tables below, whose first columns list all 4 (or 8) combinations. In each table the columns of the two sides of each equivalence agree in every row, as the reader can check column by column. Hence both sides have the same value under v, and since v was arbitrary, they are equivalent by the paraphrase of Definition 4.5.4.

Parts (a), (b), (c), and (e):

φ	ψ	$\varphi \to \psi$	$\neg\psi \to \neg\varphi$	$\neg\varphi \vee \psi$	$(\varphi \to \psi) \wedge (\psi \to \varphi)$
T	T	T	T	T	T
T	F	F	F	F	F
F	T	T	T	T	F
F	F	T	T	T	T

The first three formula columns agree, which gives (a) and (c). The last agrees with the column of $\varphi \leftrightarrow \psi$ in the table of Definition 4.4.2 (T exactly when $a = b$), which gives (e).

φ	ψ	$\neg(\varphi \wedge \psi)$	$\neg\varphi \vee \neg\psi$	$\neg(\varphi \vee \psi)$	$\neg\varphi \wedge \neg\psi$
T	T	F	F	F	F
T	F	T	T	F	F
F	T	T	T	F	F
F	F	T	T	T	T

This gives (b). For (d): if $a = \mathrm{T}$, then $v(\neg\varphi) = \mathrm{F}$ and $v(\neg\neg\varphi) = \mathrm{T}$; if $a = \mathrm{F}$, then $v(\neg\varphi) = \mathrm{T}$ and $v(\neg\neg\varphi) = \mathrm{F}$; in both cases $v(\neg\neg\varphi) = a$. For (f): clause (V3) makes $v(\varphi \wedge \psi) = \mathrm{T}$ exactly when $a = b = \mathrm{T}$, a condition that does not change when a and b are swapped, so $v(\varphi \wedge \psi) = v(\psi \wedge \varphi)$; likewise (V4) for $\vee$.

Parts (g) and (h):

φ	ψ	χ	$\varphi \wedge (\psi \vee \chi)$	$(\varphi \wedge \psi) \vee (\varphi \wedge \chi)$	$\varphi \vee (\psi \wedge \chi)$	$(\varphi \vee \psi) \wedge (\varphi \vee \chi)$
T	T	T	T	T	T	T
T	T	F	T	T	T	T
T	F	T	T	T	T	T
T	F	F	F	F	T	T
F	T	T	F	F	T	T
F	T	F	F	F	F	F
F	F	T	F	F	F	F
F	F	F	F	F	F	F

φ	ψ	χ	$(\varphi \wedge \psi) \wedge \chi$	$\varphi \wedge (\psi \wedge \chi)$	$(\varphi \vee \psi) \vee \chi$	$\varphi \vee (\psi \vee \chi)$
T	T	T	T	T	T	T
T	T	F	F	F	T	T
T	F	T	F	F	T	T
T	F	F	F	F	T	T
F	T	T	F	F	T	T
F	T	F	F	F	T	T
F	F	T	F	F	T	T
F	F	F	F	F	F	F

In each of the two tables the first two formula columns agree, and so do the last two. This proves all parts.

■

Chapters 1 and 2 derived many of these laws as pairs of sequents, some directions with derivations that use RAA or LEM (Proposition 2.5.5 (d), Proposition 2.5.1, Proposition 2.5.6) and others without. That some of them cannot do without a classical rule, Chapter 2 stated without proof (for De Morgan (d), for instance); for $\neg\neg P \vdash P$, Section 4.7 proves it in an optional result. The tables do not see that difference: they see only that the two sides agree in every row. Section 4.8 shows that every such agreement can be turned into derivations, with RAA among the rules.

Long conjunctions and disjunctions. Associativity lets us stop worrying about grouping. We write $\varphi_1 \wedge \varphi_2 \wedge \cdots \wedge \varphi_n$ for the conjunction grouped to the right, $\varphi_1 \wedge (\varphi_2 \wedge (\cdots \wedge \varphi_n))$, and likewise for $\vee$; for $n = 1$ it is φ_1 itself. Then, for every valuation v and every $n \geq 1$,

$$v(\varphi_1 \wedge \cdots \wedge \varphi_n) = \mathrm{T} \iff v(\varphi_i) = \mathrm{T} \text{ for every } i \text{ with } 1 \leq i \leq n,$$
$$v(\varphi_1 \vee \cdots \vee \varphi_n) = \mathrm{T} \iff v(\varphi_i) = \mathrm{T} \text{ for some } i \text{ with } 1 \leq i \leq n.$$

Second reading · *Why long conjunctions are evaluated term by term*

For the conjunction, we prove by Axiom 4.2.1 on n, started at 1, that the first line holds for **every** list of n formulas. For $n = 1$ both sides say $v(\varphi_1) = \mathrm{T}$. If it holds for every list of n formulas, take a list of $n+1$. The conjunction is $\varphi_1 \wedge \theta$, where $\theta = \varphi_2 \wedge \cdots \wedge \varphi_{n+1}$ has n terms. By (V3), $v(\varphi_1 \wedge \theta) = \mathrm{T}$ exactly when $v(\varphi_1) = \mathrm{T}$ and $v(\theta) = \mathrm{T}$, and by the hypothesis for n, $v(\theta) = \mathrm{T}$ exactly when $v(\varphi_i) = \mathrm{T}$ for $2 \leq i \leq n+1$. Together: exactly when every $v(\varphi_i) = \mathrm{T}$. The disjunction is the same with (V4) and "some".

The same holds for any grouping. Call a formula built from $\varphi_1, \ldots, \varphi_n$, in this order, using $\wedge$ alone, a *grouping* of the list. For $n \geq 2$ a grouping is $\theta_1 \wedge \theta_2$, where θ_1 is a grouping of $\varphi_1, \ldots, \varphi_k$ and θ_2 a grouping of $\varphi_{k+1}, \ldots, \varphi_n$, for some k with $1 \leq k < n$. Each part has fewer terms, so by (V3) and Theorem 4.2.2 on n, a grouping is true under v exactly when every φ_i is. Likewise for $\vee$ with (V4).

Two consequences follow at once, since the right-hand sides do not care about order, and by the fold every grouping has the same right-hand side. Reordering the terms of a long conjunction, or grouping them differently, gives an equivalent formula; likewise for disjunctions. And De Morgan's laws extend to many terms:

$$\neg(\varphi_1 \wedge \cdots \wedge \varphi_n) \equiv \neg\varphi_1 \vee \cdots \vee \neg\varphi_n, \qquad \neg(\varphi_1 \vee \cdots \vee \varphi_n) \equiv \neg\varphi_1 \wedge \cdots \wedge \neg\varphi_n,$$

since, for instance, both sides of the first are true exactly when some φ_i is false.

Quick check. Is $p \to (q \to r) \equiv (p \wedge q) \to r$? And is $p \to (q \to r) \equiv (p \to q) \to r$?

Solution. The first holds. Each side is false in exactly one row, $p = q = \mathrm{T}$, $r = \mathrm{F}$: the left needs p true and $q \to r$ false, the right needs $p \wedge q$ true and r false. (Chapter 1 derived the two from each other, Proposition 1.4.2.) The second fails: by Example (One grouping entails the other, not conversely), the row $p = r = \mathrm{F}$ makes $p \to (q \to r)$ true and $(p \to q) \to r$ false.

Remark. **Three ways to change a formula.** The chapter has three operations that put one thing in place of another. Their names are close, and what each one keeps is different.

Operation	What is put in	Where	What it keeps	Example
replace a part	one occurrence of a subformula, by a formula equivalent to it	below, "Replacing a part by an equivalent part"	the truth table of the whole formula	$\neg\neg r$ by r inside $p \to (q \vee \neg\neg r)$
substitute for atoms	formulas for chosen atoms, at **every** occurrence, all at once	below, "Substituting formulas for atoms"	tautologies stay tautologies, valid arguments stay valid	$q \wedge r$ for p turns $p \vee \neg p$ into $(q \wedge r) \vee \neg(q \wedge r)$
substitute in a derivation	formulas for the letters of a derivation, at every occurrence	Section 4.1, last subsection	being a derivation	$p_1 \wedge p_2$ for P in the derivation of Proposition 1.4.1

Quick check. Which of the three operations turns $p \vee \neg p$ into $(q \to r) \vee \neg(q \to r)$, and what does it keep?

Solution. Substitution of $q \to r$ for the atom p, at both occurrences. It keeps tautologies, so the new formula is a tautology too. It is not a replacement of a part: $q \to r$ is not equivalent to p.

4.5.4 Relations between two formulas

Equivalent formulas agree in every row. Three other relations say how two formulas can fail to agree, and they have old names.

Definition 4.5.5. Contradictory, Contrary and Subcontrary Formulas

Formulas φ and ψ are **contradictory** if every row makes exactly one of them true; **contrary** if no row makes both true but some row makes both false; **subcontrary** if no row makes both false but some row makes both true.

In words: contradictory formulas are each other's negation in value, $\varphi \equiv \neg\psi$, since "exactly one true" says that $v(\varphi) = \mathrm{T}$ iff $v(\psi) = \mathrm{F}$, that is, iff $v(\neg\psi) = \mathrm{T}$. Contrary formulas cannot both hold but can both fail; subcontrary formulas cannot both fail but can both hold. A row here is a valuation, or equally, by Lemma 4.4.1, a row of a table on the atoms of both formulas.

Pair	Relation	Why
$p \to q$, $p \wedge \neg q$	contradictory	each is F exactly where the other is T: $p \wedge \neg q$ is T only at TF, the one row where $p \to q$ is F
$p \wedge q$, $\neg p$	contrary	both false at $p = \mathrm{T}$, $q = \mathrm{F}$; both true would need p and $\neg p$ true
$p \vee q$, $\neg p \vee q$	subcontrary	both true at $p = \mathrm{F}$, $q = \mathrm{T}$; both false would need p and $\neg p$ false

Quick check. Classify the pair $p \leftrightarrow q$, $p \leftrightarrow \neg q$.

Solution. Contradictory. In each row the two sides of the first agree exactly when the two sides of the second disagree, since $\neg q$ has the other value from q. So exactly one of the two is true in each row.

4.5.5 Replacing a part by an equivalent part

In algebra, if $a = b$, then a may be replaced by b inside any expression. Semantic equivalence behaves in the same way, because the value of a formula depends only on the values of its parts. This is what makes chains of equivalences, one standard equivalence applied deep inside a formula at a time, a legitimate way of computing.

Recall from Section 4.1 that the subformulas of χ are the formulas at the nodes of its parse tree (Definition 4.1.3). An **occurrence** of φ in χ is a node whose subtree is a parse tree of φ; **replacing** that occurrence by φ' means putting a parse tree of φ' in place of that subtree, which gives the parse tree of a new formula χ'.

Lemma 4.5.2. Replacement

Let $\varphi \equiv \varphi'$. If χ' is obtained from a formula χ by replacing one occurrence of φ in χ by φ', then $\chi' \equiv \chi$.

Idea. Induct on the shape of χ. Either the occurrence is all of χ, and then χ is φ and χ' is φ', equivalent by hypothesis; or it sits inside one immediate part of χ, the other part is untouched, and the value of χ is computed from the values of its parts, one of which has not changed and one of which has not changed its value.

The proof, by structural induction with exactly that case split, is in the fold below.

Second reading · *The proof of the replacement lemma*

Proof. We apply Theorem 4.3.1 to the property $C(\chi)$: *for every formula χ' obtained from χ by replacing one occurrence of φ by φ', and every valuation v, $v(\chi') = v(\chi)$.* By the paraphrase of Definition 4.5.4, $C(\chi)$ says that each such χ' is equivalent to χ.

In every case, the replaced occurrence sits either at the top node of the tree of χ or at a node below it; in the second case the node lies in the subtree of exactly one immediate part of χ, since those subtrees are the parse trees of the immediate parts (Definition 4.1.2). If the occurrence is the root, then $\chi = \varphi$ and $\chi' = \varphi'$, and $v(\chi') = v(\chi)$ because $\varphi \equiv \varphi'$. It remains to treat an occurrence below the root.

Atoms and $\bot$. The parse tree is one node, so there is no node below the root.

Negation. Let $\chi = \neg\theta$ and assume $C(\theta)$. An occurrence below the root lies in θ, so $\chi' = \neg\theta'$, where θ' is obtained from θ by replacing that occurrence. By $C(\theta)$, $v(\theta') = v(\theta)$, and (V2) gives $v(\neg\theta') = v(\neg\theta)$.

Binary connectives. Let $\chi = \theta_1 \circ \theta_2$, with $\circ$ one of $\wedge, \vee, \to, \leftrightarrow$, and assume $C(\theta_1)$ and $C(\theta_2)$. If the occurrence lies in θ_1, then $\chi' = \theta_1' \circ \theta_2$, with θ_1' obtained from θ_1 by the replacement, and $v(\theta_1') = v(\theta_1)$ by $C(\theta_1)$. The clause of Definition 4.4.2 for $\circ$ computes $v(\chi')$ from $v(\theta_1')$ and $v(\theta_2)$, and $v(\chi)$ from $v(\theta_1)$ and $v(\theta_2)$, by the same rule, so $v(\chi') = v(\chi)$. An occurrence in θ_2 is the same with the roles of the parts exchanged.

This covers every clause of Definition 4.1.1, so $C(\chi)$ holds for every χ, which proves the lemma.

■

Replacing several occurrences, one after another, keeps equivalence at each step, and $\equiv$ is transitive, so the final formula is equivalent to the first. For instance, with (b), (d), and then (c) of Proposition 4.5.1, each applied inside the formula:

$$\begin{aligned} p \to \neg(q \wedge \neg r) &\equiv p \to (\neg q \vee \neg\neg r) \\ &\equiv p \to (\neg q \vee r) \\ &\equiv \neg p \vee (\neg q \vee r). \end{aligned}$$

The first step replaces the subformula $\neg(q \wedge \neg r)$ by an equivalent one (De Morgan with $\varphi = q$, $\psi = \neg r$); the second replaces $\neg\neg r$ by r; the third replaces the whole formula, by (c) with $\varphi = p$ and $\psi = \neg q \vee r$. Section 4.9 computes normal forms this way.

Warning. Replacement needs **equivalence**; entailment in one direction is not enough. We have $p \wedge q \models p$, but replacing $p \wedge q$ by p inside $\neg(p \wedge q)$ gives $\neg p$, and the row $p = \mathrm{T}$, $q = \mathrm{F}$ makes $\neg(p \wedge q)$ true and $\neg p$ false. A negation reverses the direction of an entailment, so one-way information cannot pass through it.

4.5.6 Substituting formulas for atoms

The letters P, Q of Chapters 1 and 2 stood for arbitrary statements, and Section 4.1 turned that into a remark about derivations: a derivation written with letters is a pattern, and putting formulas in for the letters, uniformly, gives a derivation again. The truth table of Example (The truth table of modus ponens), though, was computed for the atoms p and q. Does it tell us anything about $(r \vee s) \to \neg t$ in place of p? It should: the table only used the values of p and q, and a compound formula also has a value in each row. This subsection makes that precise.

Let $p_1, \ldots, p_n$ be distinct atoms and $\varphi_1, \ldots, \varphi_n$ formulas. The **substitution** σ of $\varphi_1, \ldots, \varphi_n$ for $p_1, \ldots, p_n$ is the function on formulas defined by recursion (Theorem 4.3.4):

$$\sigma(p_i) = \varphi_i \ (1 \le i \le n), \qquad \sigma(q) = q \ \text{ for every other atom } q, \qquad \sigma(\bot) = \bot,$$
$$\sigma(\neg\chi) = \neg\sigma(\chi), \qquad \sigma(\chi \circ \theta) = \sigma(\chi) \circ \sigma(\theta) \ \text{ for each binary } \circ\,.$$

In words: $\sigma(\chi)$ replaces **every** occurrence of each p_i in χ by φ_i, all at once, and changes nothing else. For a list $\Gamma = \gamma_1, \ldots, \gamma_m$, we write $\sigma(\Gamma)$ for the list $\sigma(\gamma_1), \ldots, \sigma(\gamma_m)$. For example, with σ putting $q \wedge r$ for p and p for q,

$$\sigma\big((p \to q) \vee \neg p\big) = \big((q \wedge r) \to p\big) \vee \neg(q \wedge r).$$

The p that σ puts in for q is not replaced again: the replacements happen simultaneously, in one pass over the parse tree.

Example. One row, two valuations

Let σ put $q \wedge r$ for p and p for q, as above, and let $v(p) = \mathrm{F}$, $v(q) = \mathrm{T}$, $v(r) = \mathrm{T}$. Compute $v(\sigma(\chi))$ for $\chi = (p \to q) \vee \neg p$. Then define w by $w(s) := v(\sigma(s))$ for every atom s, and compute $w(\chi)$.

Solution. By the display above, $\sigma(\chi) = ((q \wedge r) \to p) \vee \neg(q \wedge r)$. Under v, $q \wedge r$ is T and p is F, so $(q \wedge r) \to p$ is F; $\neg(q \wedge r)$ is F; so $v(\sigma(\chi)) = \mathrm{F}$.

Next, $w(p) = v(q \wedge r) = \mathrm{T}$, $w(q) = v(p) = \mathrm{F}$, and $w(r) = v(r) = \mathrm{T}$, since σ leaves r alone. Under w, $p \to q$ is F and $\neg p$ is F, so $w(\chi) = \mathrm{F}$. The two values agree.

The formula $\sigma(\chi)$ under v behaves exactly as χ does under w. The lemma below says that this happens for every formula and every valuation, and draws the consequences.

Lemma 4.5.3. Substitution

Let σ be the substitution of $\varphi_1, \ldots, \varphi_n$ for the distinct atoms $p_1, \ldots, p_n$.

(a) If $\models \chi$, then $\models \sigma(\chi)$.
(b) If $\Gamma \models \psi$, then $\sigma(\Gamma) \models \sigma(\psi)$.

Idea. Under a valuation v, the formula $\sigma(\chi)$ behaves exactly like χ itself under a different valuation w: the one that gives each atom p_i the value that v gives to φ_i. So every row for $\sigma(\chi)$ is, in disguise, a row for χ, and whatever holds in all rows of χ holds in all rows of $\sigma(\chi)$. The disguise is proved by induction on the shape of χ.

Proof. Let v be a valuation, and let w be the valuation with $w(p) := v(\sigma(p))$ for every atom p; so $w(p_i) = v(\varphi_i)$ and $w(q) = v(q)$ for every other atom q.

Claim. $v(\sigma(\chi)) = w(\chi)$ for every formula χ.

Proof of Claim. By Theorem 4.3.1. For an atom p, $v(\sigma(p)) = w(p)$ is the definition of w. For $\bot$, $v(\sigma(\bot)) = v(\bot) = \mathrm{F} = w(\bot)$ by (V1). For a negation, assume $v(\sigma(\chi)) = w(\chi)$; then $v(\sigma(\neg\chi)) = v(\neg\sigma(\chi))$, which by (V2) is the flip of $v(\sigma(\chi)) = w(\chi)$, and so equals $w(\neg\chi)$. For a binary connective $\circ$, assume the claim for χ and θ; then $v(\sigma(\chi \circ \theta)) = v(\sigma(\chi) \circ \sigma(\theta))$ is computed by the clause for $\circ$ from $v(\sigma(\chi)) = w(\chi)$ and $v(\sigma(\theta)) = w(\theta)$, and so equals $w(\chi \circ \theta)$, which the same clause computes from the same two values.

■

(a) Let $\models \chi$, and let v be any valuation. By the claim, $v(\sigma(\chi)) = w(\chi) = \mathrm{T}$, since χ is a tautology. So $\sigma(\chi)$ is true under every valuation.

(b) Let $\Gamma \models \psi$, and let v make every formula of $\sigma(\Gamma)$ true. For each γ in Γ, the claim gives $w(\gamma) = v(\sigma(\gamma)) = \mathrm{T}$, so w makes every formula of Γ true. By $\Gamma \models \psi$, $w(\psi) = \mathrm{T}$, and by the claim again $v(\sigma(\psi)) = w(\psi) = \mathrm{T}$. This proves $\sigma(\Gamma) \models \sigma(\psi)$.

■

So a table computed once for atoms settles infinitely many instances. From $\models p \vee \neg p$ we get $\models (q \to r) \vee \neg(q \to r)$, and every other instance of excluded middle. From $p \to q$, $p \models q$ (Example (The truth table of modus ponens)) we get $\varphi \to \psi$, $\varphi \models \psi$ for all formulas φ, ψ, which is the table's version of $\to$E as a schema. And from Example (Affirming the consequent is not valid)'s refuting row we learn only that the pattern is not valid, not that its instances fail; the lemma runs one way only, as the next warning says.

Warning. Substitution must be **uniform**: every occurrence of p_i is replaced by the same φ_i. Replacing only the first p of the tautology $p \to p$ by q gives $q \to p$, which is false in the row $q = \mathrm{T}$, $p = \mathrm{F}$. And the lemma runs one way only. An instance can be a tautology when the pattern is not: $p \vee q$ is not a tautology, but its instance with $\neg p$ put in for q is $p \vee \neg p$, which is. So a refuting row for a pattern with atoms shows that the **pattern** is not valid, not that every instance fails.

4.5.7 One row or every row

Every question of this section is about rows, and each comes in two halves: one half is settled by a single row, the other needs them all.

Question	A yes is shown by	A no is shown by	The same as: is this list unsatisfiable?
φ satisfiable?	one row making φ true	every row	no: a yes is one row, so this is the negation of "is the list φ unsatisfiable?"
φ a tautology?	every row	one row making φ false	the list $\neg\varphi$
$\Gamma \models \psi$?	every row (or working backwards from F)	one refuting row	the counterexample set $\Gamma, \neg\psi$
$\varphi \equiv \psi$?	every row	one row where they differ	the list $\neg(\varphi \leftrightarrow \psi)$, since $\varphi \equiv \psi$ exactly when $\models \varphi \leftrightarrow \psi$ (after Definition 4.5.4)
Γ satisfiable?	one row	every row	no: it is the negation of "is Γ unsatisfiable?"

So every question is one question in disguise: is some list unsatisfiable? A yes needs every row, a no needs one. Section 4.6 searches the rows of a list without writing them all down.

4.5.8 Two questions, side by side

We now have two ways of saying that ψ follows from Γ. They were built from different materials, and it is worth setting them next to each other before Sections 4.7 and 4.8 connect them.

	$\Gamma \vdash \psi$	$\Gamma \models \psi$
What it says	**there is** a derivation from premises in Γ to ψ	**every** valuation making all of Γ true makes ψ true
What it is about	syntax: lines, rules, boxes	meaning: truth values
How to show it holds	exhibit one derivation	check every row (or work backwards from F)
How to show it fails	no method yet: a failed search proves nothing	exhibit one refuting row
Defined in	Definition 1.2.3 (Section 1.2)	Definition 4.5.3 (this section)

Look at the fourth row. To show that something is **not** derivable, we would have to survey infinitely many derivations; to show that something is **not** valid, one row suffices. And to show that something **is** valid takes a finite, mechanical check, while showing it derivable may need an idea. Each column is easy exactly where the other is hard.

Sections 4.7 and 4.8 prove that the two columns always give the same answer. Section 4.7 proves **soundness**: if $\Gamma \vdash \psi$, then $\Gamma \models \psi$; so a refuting row proves that no derivation exists, which

answers the question Chapter 3 ended on. Section 4.8 proves **completeness**: if $\Gamma \models \psi$, then $\Gamma \vdash \psi$; so a successful table check guarantees that a derivation exists, even before anyone has found it.

4.5.9 Exercises

4.5.9.1 A. Check your understanding

Exercise. A1

(a) Classify each formula as a tautology, a contradiction, or satisfiable but not a tautology: $p \to (q \to p)$; $p \wedge \neg(p \vee q)$; $(p \to q) \wedge (q \to p)$; $\neg(p \to p)$; $(p \wedge q) \to (p \vee r)$.

(b) For each, decide whether it holds, and give a refuting row if it does not: $p \wedge q \models p \leftrightarrow q$; $p \leftrightarrow q \models p \wedge q$; $\neg p \to p \models p$.

(c) True or false, with a reason: if $\Gamma \models \psi$ and $\Gamma \models \neg\psi$, then Γ is unsatisfiable.

(d) True or false, with a reason: $p \vee q \models p$ or $p \vee q \models q$.

Solution. *(a)* $p \to (q \to p)$: a tautology (if p is T, the conclusion $q \to p$ is T; if p is F, the hypothesis is false). $p \wedge \neg(p \vee q)$: a contradiction (if p is T, so is $p \vee q$, and $\neg(p \vee q)$ is F). $(p \to q) \wedge (q \to p)$: satisfiable, not a tautology (T in the row TT, F in TF). $\neg(p \to p)$: a contradiction, since $p \to p$ is a tautology. $(p \wedge q) \to (p \vee r)$: a tautology (if the hypothesis is T, then p is T, so is $p \vee r$).

(b) $p \wedge q \models p \leftrightarrow q$ holds: the only row making $p \wedge q$ true has $p = q =$ T, where $p \leftrightarrow q$ is true. $p \leftrightarrow q \models p \wedge q$ fails: the row $p = q =$ F makes $p \leftrightarrow q$ true and $p \wedge q$ false. $\neg p \to p \models p$ holds: in the row $p =$ F, the premise $\neg p \to p$ has a true hypothesis and a false conclusion, so it is false; the only row making the premise true has $p =$ T.

(c) True. If a valuation made all of Γ true, it would make both ψ and $\neg\psi$ true, which (V2) forbids. So no valuation makes all of Γ true.

(d) False. The row $p =$ F, $q =$ T refutes $p \vee q \models p$, and the row $p =$ T, $q =$ F refutes $p \vee q \models q$. Each disjunct fails in some row where the disjunction holds, although in every such row one of them holds: $\models$ must work row by row with a single conclusion.

Exercise. A2: Drill

The parts run from classifying one formula, through lists and entailments, to pairs of formulas and the reduction to unsatisfiability.

(a) Tautology, contradiction or contingent: $(p \to q) \vee p$?

(b) The same for $(p \vee q) \wedge \neg p \wedge \neg q$.

(c) The same for $(p \wedge q) \to \neg p$.

(d) Is the list $p \vee q$, $\neg p$, $\neg q$ satisfiable?

(e) Does $p \to q \models \neg p \vee q$?

(f) Give a refuting row for $p \vee q \models p \wedge q$.

(g) Does $\bot \models p$?

(h) Does $p \models q \to p$?

(i) Is $\neg(p \to q) \equiv p \wedge \neg q$?
(j) Classify the pair $p \wedge q$, $\neg p \vee \neg q$: contradictory, contrary, subcontrary, or none.
(k) Classify the pair $p \wedge \neg q$, q.
(l) Which list must be unsatisfiable for $p \to q \equiv \neg q \to \neg p$ to hold?

Solution.

(a) A tautology. If p is F, the left disjunct is T; if p is T, the right one is.
(b) A contradiction: $\neg p$ and $\neg q$ true leave $p \vee q$ false.
(c) Contingent: F at $p = q =$ T, and T whenever p is F.
(d) No: a row making $\neg p$ and $\neg q$ true makes $p \vee q$ false.
(e) Yes. The two are equivalent, by Proposition 4.5.1 (c).
(f) $p =$ T, $q =$ F.
(g) Yes: no row makes $\bot$ true.
(h) Yes: in every row with p true, $q \to p$ has a true conclusion.
(i) Yes: both are T in the row TF only.
(j) Contradictory: by Proposition 4.5.1 (b) the second is equivalent to $\neg(p \wedge q)$.
(k) Contrary: never both true, since the first needs q false; both false at $p = q =$ F.
(l) The one-formula list $\neg((p \to q) \leftrightarrow (\neg q \to \neg p))$.

4.5.9.2 B. Practice

Exercise. B1: Valid or not

Decide whether each entailment holds, by a table or by working backwards from F. If it holds, also give a derivation of the corresponding sequent with $\vdash$ in place of $\models$, using the rules of Chapters 1 and 2; if it fails, give a refuting row.

(a) $p \to q,\ \neg q \models \neg p$.
(b) $p \to (q \vee r),\ \neg q \models p \to r$.
(c) $p \to q \models \neg p \to \neg q$.
(d) $p \vee q,\ p \to r,\ q \to r \models r$.

Solution. (a) Holds. Backwards from F: $\neg p$ false means $p =$ T; $\neg q$ true means $q =$ F; then $p \to q$ is F, so no row makes the premises true and the conclusion false. A derivation, with modus tollens (Proposition 2.3.1) as a derived rule:

1	$p \to q$	
2	$\neg q$	
3	$\neg p$	MT 1, 2

(b) Holds. Backwards from F: $p \to r$ false means $p =$ T, $r =$ F; $\neg q$ true means $q =$ F; then $q \vee r$ is F, and $p \to (q \vee r)$ is F, contradicting the first premise. A derivation, with disjunctive syllogism (Proposition 2.3.4), which refutes the left disjunct:

1	$p \to (q \vee r)$	
2	$\neg q$	
3	$\quad p$	asm, for $\to$I
4	$\quad q \vee r$	$\to$E 1, 3
5	$\quad r$	DS 4, 2
6	$p \to r$	$\to$I 3–5

(c) Fails. The row $p = \mathrm{F}$, $q = \mathrm{T}$ makes $p \to q$ true (false hypothesis) and $\neg p \to \neg q$ false ($\neg p$ true, $\neg q$ false). It is the pattern "denying the hypothesis" of the Quick check $p \to q,\ \neg p \models \neg q$, now in the form of an implication.

(d) Holds. Backwards from F: $r = \mathrm{F}$; then $p \to r$ true forces $p = \mathrm{F}$, and $q \to r$ true forces $q = \mathrm{F}$; then $p \vee q$ is F, a contradiction. A derivation by cases:

1	$p \vee q$	
2	$p \to r$	
3	$q \to r$	
4	$\quad p$	asm, for $\vee$E
5	$\quad r$	$\to$E 2, 4
6	$\quad q$	asm, for $\vee$E
7	$\quad r$	$\to$E 3, 6
8	r	$\vee$E 1, 4–5, 6–7

Exercise. B2: A biconditional as a disjunction

Show by a truth table that $p \leftrightarrow q \equiv (p \wedge q) \vee (\neg p \wedge \neg q)$. Which theorem of this chapter will guarantee, without any search, that each side can be derived from the other?

Solution.

p	q	$p \leftrightarrow q$	$p \wedge q$	$\neg p \wedge \neg q$	$(p \wedge q) \vee (\neg p \wedge \neg q)$
T	T	T	T	F	T
T	F	F	F	F	F
F	T	F	F	F	F
F	F	T	F	T	T

The columns of $p \leftrightarrow q$ and of the disjunction agree in every row, so the two formulas are equivalent: each side entails the other. The theorem that turns each entailment into a derivation is completeness, in Section 4.8: if $\Gamma \models \psi$, then $\Gamma \vdash \psi$. Once it is proved, it will give $p \leftrightarrow q \vdash (p \wedge q) \vee (\neg p \wedge \neg q)$ and its converse, before any derivation has been written; until then the table shows only that the two sequents are valid.

Exercise. B3: Entailment and implication

Let Γ be a list of formulas and φ, ψ formulas. Prove:

(a) $\varphi \models \psi$ if and only if $\models \varphi \to \psi$;
(b) $\Gamma, \varphi \models \psi$ if and only if $\Gamma \models \varphi \to \psi$.

Solution. We prove (b); then (a) is (b) with Γ empty, since Γ, φ is then the one-formula list φ.

($\Rightarrow$) Suppose $\Gamma, \varphi \models \psi$, and let v make every formula of Γ true. If $v(\varphi) = \mathrm{F}$, then $v(\varphi \to \psi) = \mathrm{T}$ by (V5). If $v(\varphi) = \mathrm{T}$, then v makes every formula of the list Γ, φ true, so $v(\psi) = \mathrm{T}$ by hypothesis, and again $v(\varphi \to \psi) = \mathrm{T}$. So $\Gamma \models \varphi \to \psi$.

($\Leftarrow$) Suppose $\Gamma \models \varphi \to \psi$, and let v make every formula of Γ, φ true. Then v makes every formula of Γ true, so $v(\varphi \to \psi) = \mathrm{T}$, and $v(\varphi) = \mathrm{T}$. By (V5), an implication with a true hypothesis is true only if its conclusion is, so $v(\psi) = \mathrm{T}$. So $\Gamma, \varphi \models \psi$.

Part (b) is the table's version of $\to$I (the direction $\Rightarrow$) and of $\to$E (the direction $\Leftarrow$).

Exercise. B4: Contradictory, contrary, subcontrary, or none

Classify each pair, and give the rows that show it.

(a) $p \to q$, $p \wedge \neg q$.
(b) $p \wedge q$, $\neg p \wedge q$.
(c) $p \vee q$, $\neg p \vee q$.
(d) $p \to q$, $q \to p$.
(e) $p \leftrightarrow q$, $p \wedge \neg q$.
(f) p, $p \vee q$.

Solution. (a) Contradictory: the first is F only at TF, where the second is T, and the second is T only there.

(b) Contrary: both true would need p and $\neg p$ true; both are false at $q = \mathrm{F}$.

(c) Subcontrary: both true at $p = \mathrm{F}$, $q = \mathrm{T}$; both false would need q false and both p and $\neg p$ false.

(d) Subcontrary: both true at TT; $p \to q$ is F only at TF and $q \to p$ only at FT, so they are never both false.

(e) Contrary: both true would need p and q equal (for $p \leftrightarrow q$) and unequal (for $p \wedge \neg q$); both false at $p = \mathrm{F}$, $q = \mathrm{T}$.

(f) None of the three: both true at $p = \mathrm{T}$, both false at $p = q = \mathrm{F}$. (In fact $p \models p \vee q$.)

The pair (d) is the surprise: an implication and its converse can never both fail.

Marking scheme. (6 points)

- 1 per part: the class, with a witness for each of the two questions (both true? both false?): a row where it happens, or the reason it cannot.

Fatal error (0 for the part): a class claimed with one of the two questions unanswered.

4.5.9.3 C. Going deeper

Exercise. C1: Substitution with a new connective, the last half blank, used later

Extend the formulas by a new binary connective $\mid$ ("not both", often called NAND): add to Definition 4.1.1 the clause "if φ and ψ are formulas, so is $(\varphi \mid \psi)$", and to Definition 4.4.2 the clause

(V7) $v(\varphi \mid \psi) := \mathrm{F}$ if $v(\varphi) = v(\psi) = \mathrm{T}$; otherwise T.

Assume that structural induction and recursion work for the extended formulas as before, with one more clause. A substitution σ of $\varphi_1, \ldots, \varphi_n$ for $p_1, \ldots, p_n$ is defined by the clauses in the text, with $\sigma(\chi \mid \theta) := \sigma(\chi) \mid \sigma(\theta)$ added. Prove that if χ is a tautology of the extended language, so is $\sigma(\chi)$.

Here is the first half of a proof. *Let v be a valuation, and let w be the valuation with $w(p) := v(\sigma(p))$ for every atom p. We show $v(\sigma(\chi)) = w(\chi)$ for every formula χ of the extended language, by structural induction. Atoms: $v(\sigma(p)) = w(p)$ is the definition of w.*

Name the move, then complete it.

Solution. The move is **induction on the shape of a formula**, now with eight cases: atoms, $\bot$, $\neg$, and five binary connectives. The property is $C(\chi)$: $v(\sigma(\chi)) = w(\chi)$.

$\bot$. $\sigma(\bot) = \bot$, so $v(\sigma(\bot)) = \mathrm{F} = w(\bot)$ by (V1).

Negation. Assume $C(\chi)$. Then $v(\sigma(\neg\chi)) = v(\neg\sigma(\chi))$, which by (V2) is T exactly when $v(\sigma(\chi)) = \mathrm{F}$, that is, when $w(\chi) = \mathrm{F}$, that is, when $w(\neg\chi) = \mathrm{T}$.

The four old binary connectives. Assume $C(\chi)$ and $C(\theta)$, and let $\circ$ be one of $\wedge, \vee, \to, \leftrightarrow$. Then $v(\sigma(\chi \circ \theta)) = v(\sigma(\chi) \circ \sigma(\theta))$ is computed by the clause for $\circ$ from the two values $v(\sigma(\chi)) = w(\chi)$ and $v(\sigma(\theta)) = w(\theta)$; the same clause computes $w(\chi \circ \theta)$ from the same two values. So they are equal.

The new connective. Assume $C(\chi)$ and $C(\theta)$. By (V7), $v(\sigma(\chi|\theta)) = v(\sigma(\chi)|\sigma(\theta))$ is F exactly when $v(\sigma(\chi)) = v(\sigma(\theta)) = \mathrm{T}$, that is, when $w(\chi) = w(\theta) = \mathrm{T}$, that is, when $w(\chi \mid \theta) = \mathrm{F}$.

This covers every clause, so $C(\chi)$ holds for every formula. Now let χ be a tautology and v any valuation: $v(\sigma(\chi)) = w(\chi) = \mathrm{T}$. So $\sigma(\chi)$ is a tautology.

Only one case is new. The proof of Lemma 4.5.3 never looked at what the clauses say, only that each computes the value of a compound from the values of its parts; any connective defined by a table passes.

Marking scheme. (10 points)

- 1: names the move and states the property $v(\sigma(\chi)) = w(\chi)$ for all χ.
- 1: the $\bot$ case.
- 2: the $\neg$ case, from the hypothesis for χ.

- 2: the old binary cases, using the hypotheses for both parts.
- 3: the | case, with (V7) and both hypotheses.
- 1: the conclusion for tautologies.

Fatal errors (0 for the step): omitting the | case; using w where v is meant in the conclusion.

Exercise. C2: Is equivalence preserved by substitution?

Let σ be a substitution and suppose $\varphi \equiv \psi$. Is $\sigma(\varphi) \equiv \sigma(\psi)$? Prove it, or give a counterexample. Then suppose that De Morgan's law (b) of Proposition 4.5.1 had been checked only for atoms, $\neg(p \wedge q) \equiv \neg p \vee \neg q$, by its four-row table. Use your answer to deduce from this, in one line, that $\neg((r \to s) \wedge \neg t) \equiv \neg(r \to s) \vee \neg\neg t$.

Solution. Yes. By the paraphrase of Definition 4.5.4, $\varphi \equiv \psi$ means $\models \varphi \leftrightarrow \psi$. By Lemma 4.5.3 (a), $\models \sigma(\varphi \leftrightarrow \psi)$, and by the definition of σ, $\sigma(\varphi \leftrightarrow \psi) = \sigma(\varphi) \leftrightarrow \sigma(\psi)$. So $\models \sigma(\varphi) \leftrightarrow \sigma(\psi)$, that is, $\sigma(\varphi) \equiv \sigma(\psi)$. (Alternatively, apply Lemma 4.5.3 (b) to $\varphi \models \psi$ and to $\psi \models \varphi$.)

For the instance: the substitution of $r \to s$ for p and $\neg t$ for q turns the two sides of $\neg(p \wedge q) \equiv \neg p \vee \neg q$ into $\neg((r \to s) \wedge \neg t)$ and $\neg(r \to s) \vee \neg\neg t$, so these are equivalent by the first part. Checking a law for atoms, by one small table, is therefore enough.

A tempting counterexample fails for a reason worth seeing. With σ putting $\neg p$ for q, the nonequivalent formulas p and $\neg q$ become the equivalent p and $\neg\neg p$: substitution can create equivalences. It cannot destroy one, which is what the question asks.

Marking scheme. (10 points)

- 3: reduces $\equiv$ to a tautology $\models \varphi \leftrightarrow \psi$, or to two entailments, citing the definition.
- 3: applies Lemma 4.5.3 correctly.
- 2: in the $\leftrightarrow$ route, computes $\sigma(\varphi \leftrightarrow \psi) = \sigma(\varphi) \leftrightarrow \sigma(\psi)$ from the definition of σ; in the entailment route, applies (b) in both directions.
- 2: the instance, with the substitution named.

Fatal errors (0 for the step): answering "no" by a pair that is not equivalent before the substitution; using a nonuniform replacement.

Exercise. C3: The rows of implication are forced

Pretend we do not know the table of $\to$. Say only that there is **some** table: a rule t giving a truth value $t(a, b)$ for each pair of truth values a, b, with

$$v(\varphi \to \psi) = t\big(v(\varphi), v(\psi)\big)$$

for every valuation v and all formulas φ, ψ. (The other connectives keep their clauses.) Suppose the two rules for implication preserve truth in the following sense.

- (E) For every valuation v and all formulas φ, ψ: if $v(\varphi \to \psi) = \mathrm{T}$ and $v(\varphi) = \mathrm{T}$, then $v(\psi) = \mathrm{T}$.
-

(I) For every list of formulas Γ and all formulas φ, ψ: if $\Gamma, \varphi \models \psi$, then $\Gamma \models \varphi \to \psi$, where $\models$ is read with the table t for $\to$.

Prove that t is the table of (V5): $t(\mathrm{T}, \mathrm{F}) = \mathrm{F}$, and $t(a, b) = \mathrm{T}$ for the other three pairs.

Solution (Hint 1). Each of the four values of t is one instance of (E) or of (I), with atoms for φ and ψ and a well-chosen valuation.

Solution (Hint 2). For $t(\mathrm{T}, \mathrm{T})$ and $t(\mathrm{F}, \mathrm{F})$, try (I) with Γ empty and $\varphi = \psi = p$.

Solution. Let p and q be distinct atoms.

$t(\mathrm{T}, \mathrm{F}) = \mathrm{F}$. Take a valuation v with $v(p) = \mathrm{T}$, $v(q) = \mathrm{F}$. If $t(\mathrm{T}, \mathrm{F})$ were T, then $v(p \to q) = t(\mathrm{T}, \mathrm{F}) = \mathrm{T}$ and $v(p) = \mathrm{T}$, so (E) would give $v(q) = \mathrm{T}$, which is false. So $t(\mathrm{T}, \mathrm{F}) = \mathrm{F}$.

$t(\mathrm{T}, \mathrm{T}) = t(\mathrm{F}, \mathrm{F}) = \mathrm{T}$. Apply (I) with Γ empty and $\varphi = \psi = p$. Its hypothesis $p \models p$ holds. So every valuation makes $p \to p$ true. A valuation with $v(p) = \mathrm{T}$ gives $t(\mathrm{T}, \mathrm{T}) = v(p \to p) = \mathrm{T}$, and one with $v(p) = \mathrm{F}$ gives $t(\mathrm{F}, \mathrm{F}) = \mathrm{T}$.

$t(\mathrm{F}, \mathrm{T}) = \mathrm{T}$. Apply (I) with Γ the list q, $\varphi = p$ and $\psi = q$. $q, p \models q$ holds, so $q \models p \to q$. A valuation with $v(p) = \mathrm{F}$, $v(q) = \mathrm{T}$ makes q true, so $t(\mathrm{F}, \mathrm{T}) = v(p \to q) = \mathrm{T}$.

No step used a value of t that had not yet been found, since the formulas in Γ, φ and ψ contain no $\to$. So the only table of $\to$ under which both rules preserve truth is the table of (V5). Conditions (E) and (I) are what Section 4.7 proves about the rules $\to$E and $\to$I; the exercise shows that nothing but the standard table could have passed.

Marking scheme. (10 points)

- 3: the row (T, F) from (E), with a valuation making p true and q false, and the argument by contradiction.
- 3: the rows (T, T) and (F, F) from (I), with the hypothesis of (I) checked, and a valuation for each value of p.
- 3: the row (F, T) from (I) with a non-empty Γ (or any other correct instance), hypothesis checked.
- 1: says why the argument is not circular (no $\to$ in the formulas used).

Fatal errors (0 for the step): using the standard table of $\to$ to justify a value of t; an instance of (I) whose hypothesis is not checked.

Exercise. C4: One question for all

Prove:

(a) $\varphi \equiv \psi$ if and only if the list $\neg(\varphi \leftrightarrow \psi)$ is unsatisfiable;

(b) for a list Γ: Γ is unsatisfiable if and only if $\Gamma \models \chi$ for every formula χ, if and only if $\Gamma \models \bot$;

(c) φ and ψ are contrary if and only if $\varphi \models \neg\psi$ and the list $\neg\varphi, \neg\psi$ is satisfiable;

(d) φ and ψ are contradictory if and only if $\varphi \equiv \neg\psi$.

Solution (Hint 1). Write out what each side says about rows: "every row ..." or "some row ...". Then compare the two sentences.

Solution. *(a)* $\varphi \equiv \psi$ iff $\models \varphi \leftrightarrow \psi$, by the display after Definition 4.5.4, iff $\neg(\varphi \leftrightarrow \psi)$ is unsatisfiable, by the equivalence between tautologies and unsatisfiable negations, applied to $\varphi \leftrightarrow \psi$.

(b) Three implications in a cycle. If Γ is unsatisfiable, then $\Gamma \models \chi$ for every χ, since there is no row to check. If $\Gamma \models \chi$ for every χ, then in particular $\Gamma \models \bot$. If $\Gamma \models \bot$, then a row making every member of Γ true would make $\bot$ true, which (V1) forbids; so no row makes Γ true, and Γ is unsatisfiable.

(c) "No row makes both true" says that every row making φ true makes ψ false, that is, makes $\neg\psi$ true: $\varphi \models \neg\psi$. "Some row makes both false" says that some row makes $\neg\varphi$ and $\neg\psi$ true: the list $\neg\varphi, \neg\psi$ is satisfiable.

(d) "Exactly one is true in every row" says that for every valuation v, $v(\varphi) = \mathrm{T}$ iff $v(\psi) = \mathrm{F}$, iff $v(\neg\psi) = \mathrm{T}$ by (V2). That is $v(\varphi) = v(\neg\psi)$ for every v, which is $\varphi \equiv \neg\psi$ by the paraphrase of Definition 4.5.4.

Marking scheme. (10 points)

- 2: (a), naming both facts used.
- 3: (b), the three implications in a cycle.
- 3: (c), both halves.
- 2: (d).

Fatal error (0 for the step): in (b), "$\Gamma \models \bot$, so Γ is false" (a list is not true or false; rows make its members true or false).

4.6 Hunting for a Refuting Row

Section 4.5 reduced every question about rows to one: is this list unsatisfiable? A truth table answers it by brute force, and with n atoms it has 2^n rows. Working backwards from F (Example (Working backwards from F)) is faster: assume the conclusion false and follow what is forced. But where a value leaves a choice, as a false conjunction or a true disjunction does, it only says "split into cases", and keeps no record of the cases.

This section turns working backwards into a method that keeps that record: take the forced steps first, and when nothing is forced, split into cases and follow every case. On paper the method is a tree. A branch that closes is a case with no row; a branch that stays open hands us a row. The answers are about truth, "valid" or "here is a refuting row"; what they say about derivations is the next section's business.

4.6.1 The rules of the search

To decide $\Gamma \models \psi$, we look for one row that makes every formula of the counterexample set $\Gamma, \neg\psi$ true; the entailment holds exactly when there is none. To decide whether a list Δ is satisfiable, we look for one row that makes all of Δ true. Requirements on the row are kept as formulas: "φ must be true" is written φ, and "φ must be false" is written $\neg\varphi$.

Quick check. From the tables of (V3)–(V5), counting each way as a condition on one or both parts: in how many ways can $\varphi \wedge \psi$ be true? In how many ways can it be false? In how many ways can $\varphi \to \psi$ be false? In how many ways can it be true?

Solution. $\varphi \wedge \psi$ is true in one way (both parts true) and false in two (φ false, or ψ false; the row with both false is covered by either). $\varphi \to \psi$ is false in one way (φ true and ψ false) and true in two (φ false, or ψ true).

Each requirement on a compound formula is traded for requirements on its parts, read off its clause. These are the **search rules**; first the requirements met in only one way.

On the branch	Its requirement, read off	Adds
$\varphi \wedge \psi$	(V3): both parts true	φ, ψ
$\neg(\varphi \vee \psi)$	(V4): both parts false	$\neg\varphi$, $\neg\psi$
$\neg(\varphi \to \psi)$	(V5): hypothesis true, conclusion false	φ, $\neg\psi$
$\neg\neg\varphi$	(V2) twice: φ true	φ

Then the requirements met in two ways; each splits the search into two cases.

On the branch	Its requirement, read off	Adds
$\varphi \vee \psi$	(V4): some part true	φ **or** ψ
$\neg(\varphi \wedge \psi)$	(V3): some part false	$\neg\varphi$ **or** $\neg\psi$
$\varphi \to \psi$	(V5): hypothesis false, or conclusion true	$\neg\varphi$ **or** ψ
$\varphi \leftrightarrow \psi$	(V6): the parts agree	φ, ψ **or** $\neg\varphi, \neg\psi$
$\neg(\varphi \leftrightarrow \psi)$	(V6): the parts differ	$\varphi, \neg\psi$ **or** $\neg\varphi, \psi$

An atom, a negated atom, and $\neg\bot$ (true in every row) need no rule. A list of requirements that holds some formula together with its negation, or holds $\bot$, describes no row at all: it is **closed**. The one-way rules are the forced steps of working backwards from F, with $\neg\neg$ added; the two-way rules are where it split into cases, with $\leftrightarrow$ added.

The two ways of a split need not exclude each other (a row may make both φ and ψ true); they need only **cover** every row that makes the formula true.

4.6.2 The search, drawn as a tree

Definition 4.6.1. Search Tree

Let Δ be a finite list of formulas. A **search tree** for Δ is built in stages. It begins as one node, the **root**, carrying the formulas of Δ. At each stage, choose an open branch B and a formula χ on B that is not an atom, a negated atom, $\bot$ or $\neg\bot$, and that is not yet **used** on B; then, below the last node of B, add χ's rule: one new node with the formulas of its one way, or two new nodes side by side with the formulas of its two ways. On each new branch, χ (and every formula equal to it) is now used: a formula is **used** on a branch exactly when the branch passes through a node added by its rule, so it stays used on every branch extending that one. Being used is a property of a formula **on a branch**: a formula at the root may be used on one branch and not on another.

A **branch** is a path from the root down to a node with nothing below it, and its formulas are those on its nodes. A branch is **closed** if it holds some formula together with its negation, or holds $\bot$; otherwise it is **open**. An open branch is **finished** if every formula on it, other than an atom, a negated atom or $\neg\bot$, is used on it. A search tree is **finished** if each of its branches is closed or finished.

In words: a search tree is a record of a backwards computation with every choice followed. A closed branch is a case with no row; a finished open branch is a case that has survived every requirement.

We draw search trees with three marks. A closed branch ends in $\times$, and a finished open branch in $\circ$. A formula is ticked ($\checkmark$) once its rule has been applied below it on some branch; the nodes its rule added show on which branches it is used. The formulas that a one-way rule adds are written straight under the stack they extend, with no new edge.

Example. A tree that closes

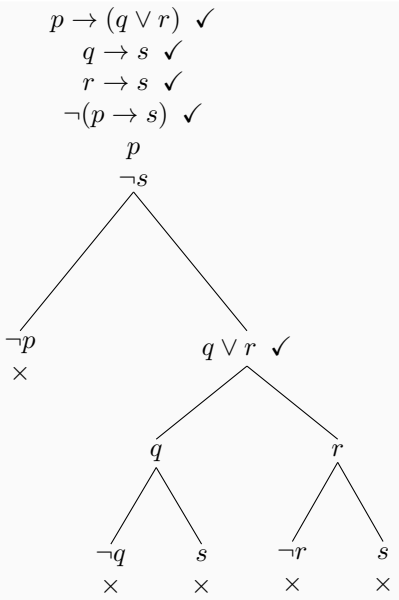
Decide whether $p \to (q \vee r),\ q \to s,\ r \to s \models p \to s$.

Solution. This is the list of Section 3.2's exit check (Exercise (C1: Exit check: the counterexample set, used later)), which derived the claim; now a search shows it valid.

The root carries the counterexample set: the three premises and $\neg(p \to s)$.

- Stage 1, the one-way rule for $\neg(p \to s)$: add p, $\neg s$.
- Stage 2, $p \to (q \vee r)$: $\neg p$ or $q \vee r$. The branch $\neg p$ closes against p.
- Stage 3, $q \vee r$: q or r.
- Stage 4, on the branch q: $q \to s$ gives $\neg q$, which closes against q, or s, which closes against $\neg s$.
- Stage 5, on the branch r: $r \to s$ gives $\neg r$ or s, both closed.

(The branch r closes without $q \to s$, so that formula is never used there.) Every branch is closed: no row makes the counterexample set true, and the entailment holds.



A tick marks a formula whose rule was applied below it: $q \to s$ was used on the branch through q, and $r \to s$ on the branch through r. Each branch closed before the other formula was needed.

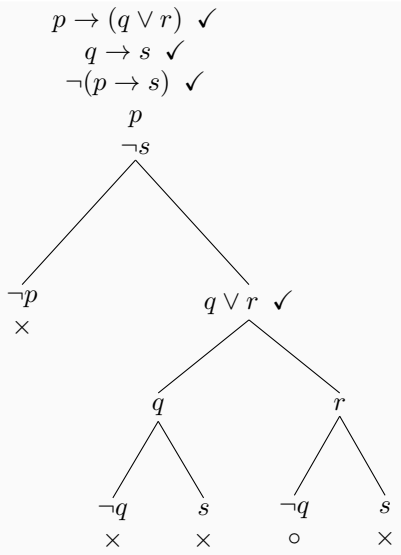
Five closed branches replace the $2^4 = 16$ rows of a table. Read as a proof by cases, the tree says: p is forced, so $q \vee r$; in the case q, s follows, against $\neg s$; in the case r, likewise.

Example. A branch that stays open

Decide whether $p \to (q \vee r),\ q \to s \models p \to s$.

Solution. The first three stages are the same as before. On the branch q, the rule of $q \to s$ closes both sides, as before.

On the branch r, at a stage of its own, $q \to s$ gives $\neg q$ or s. (A formula used on one branch may be used again on another.) The side s closes against $\neg s$; the side $\neg q$ does not. That branch carries $p \to (q \vee r)$ (used), $q \to s$ (used), $\neg(p \to s)$ (used), p, $\neg s$, $q \vee r$ (used), r and $\neg q$: it is open and finished.



Its row: the atoms on it, p and r, are true; the atoms that stand negated on it, q and s, are false. Check against the root. $p \to (q \vee r)$ is T → T, true. $q \to s$ has a false hypothesis, so it is true. $p \to s$ is T → F, false. So this row refutes the claim.

The premise $r \to s$ of the first example is exactly what closed this branch. Dropping it breaks the claim, as replacing it by $r \to t$ did in part (d) of Exercise (C1: Exit check: the counterexample set, used later), and now the search finds the row. The search may stop at the first finished open branch: one refuting row settles the question.

Quick check. On the open branch of the second tree, check that its row makes every formula on the branch true, the used ones included.

Solution. p, r, $\neg q$ and $\neg s$ are true; $q \vee r$ is true, by r; $p \to (q \vee r)$ is true, its conclusion being true; $q \to s$ is true, its hypothesis being false; and $\neg(p \to s)$ is true, since $p \to s$ is false.

4.6.3 What a closed tree proves

Each rule passes a row on. If a row makes every formula on a branch true and a rule is applied there, the row makes true all the formulas of at least one new node, since each rule was read off its table. So a row making the root true runs down some branch to its end with every formula true there. A closed branch holds φ and $\neg\varphi$, or $\bot$, which no row makes true.

An open finished branch hands over a row. Make the atoms that stand on it true and all other atoms false. A negated atom on it is then true, since its atom is not on the open branch; every other formula on it is used, so the branch carries the shorter formulas of one of its ways, and it is true by its clause once they are. From the shortest up, every formula on the branch is true.

Proposition 4.6.1. The Search Decides

Let Δ be a finite list of formulas.

(a) Every search tree for Δ can be continued until it is finished, and it has only finitely many stages.

(b) If every branch of a search tree for Δ is closed, then Δ is unsatisfiable.

(c) If a search tree for Δ has an open finished branch B, then the valuation v_B that makes the atoms on B true and every other atom false makes every formula on B true, in particular every formula of Δ.

So, for a list Γ and a formula ψ: a search tree for $\Gamma, \neg\psi$, continued as in (a), either closes, and then $\Gamma \models \psi$, or has an open finished branch, whose row refutes $\Gamma \models \psi$.

In words: the search always ends; a closed tree proves unsatisfiability, and an open finished branch exhibits a satisfying row. Since (b) and (c) cannot both apply to finished trees of one list, every order of moves gives the same verdict.

Idea.

(a) Every new formula is shorter than the one whose rule added it, and is a subformula of a root formula or the negation of one. (b) and (c) are the two observations, made exact by induction (on the stages for (b), on the length of a formula for (c)). The proof is in the fold below.

Second reading · *The proof: why the search decides*

Proof.

(a) Let N be the number of formulas that are subformulas of formulas of Δ, or negations of such. Every formula ever added to the tree is one of these. Each rule adds immediate parts of the used formula, or negations of them, or, for $\neg\neg\varphi$, the formula φ; and if the used formula is a subformula of a root formula or the negation of one, then so are the formulas it adds. Once a formula is used on a branch, it and every formula equal to it stay used on every branch below, so a stage below a given node never uses a formula already used above it. Hence each of the N formulas is used at most once along a branch, and a branch has at most N stages below the root. A stage adds at most two nodes, below one branch, so a branch splits at most N times, and the tree never has more than 2^N branches, each with at most $N + 1$ nodes. Each stage adds a node, so there are only finitely many stages. When no stage is possible, no open branch carries an unused formula other than an atom, a negated atom, $\bot$ or $\neg\bot$; since $\bot$ is on no open branch, every branch is closed or finished.

(b) *Claim:* at every stage, if a valuation v makes every formula of Δ true, then some branch has all its formulas true under v. At the start this is the root. Suppose it holds before a stage, and the stage applies χ's rule below the chosen branch B'. A branch B all of whose formulas are true under v is either another branch, which the stage leaves unchanged, or B' itself. In the second case v makes χ true, since χ is on B', and by the clause of Definition 4.4.2 from which the rule was read, v makes true every formula of one new node. One rule at a time:

- $\varphi \wedge \psi$ true: φ and ψ true, by (V3).
- $\neg(\varphi \vee \psi)$ true: $\varphi \vee \psi$ false, so $\neg\varphi$ and $\neg\psi$ true, by (V2) and (V4).
- $\neg(\varphi \to \psi)$ true: $\varphi \to \psi$ false, so φ and $\neg\psi$ true, by (V2) and (V5).

- $\neg\neg\varphi$ true: $\neg\varphi$ false, so φ true, by (V2) twice.
- $\varphi \vee \psi$ true: φ true or ψ true, by (V4).
- $\neg(\varphi \wedge \psi)$ true: $\varphi \wedge \psi$ false, so $\neg\varphi$ true or $\neg\psi$ true, by (V2) and (V3).
- $\varphi \to \psi$ true: φ false or ψ true, so $\neg\varphi$ true or ψ true, by (V5) and (V2).
- $\varphi \leftrightarrow \psi$ true: the parts agree, so both are true, or $\neg\varphi$ and $\neg\psi$ are both true, by (V6) and (V2).
- $\neg(\varphi \leftrightarrow \psi)$ true: the parts differ, so φ and $\neg\psi$ are true, or $\neg\varphi$ and ψ are, by (V2) and (V6).

So the new branch through that node has all its formulas true under v, and the claim holds after the stage. By Axiom 4.2.1 on the number of stages, it holds for the tree at hand. A closed branch holds φ and $\neg\varphi$, or $\bot$, and v cannot make all of these true, by (V2) and (V1). So if every branch is closed, no valuation makes every formula of Δ true: Δ is unsatisfiable.

(c) Let B be open and finished. We prove that $v_B(\chi) = \mathrm{T}$ for every formula χ on B, by Theorem 4.2.2 on the length of χ: assume it for every formula on B shorter than χ.

- An atom p on B is true, by the definition of v_B.
- $\bot$ is not on B, since B is open.
- $\neg\bot$ is true, by (V1) and (V2).
- A negated atom $\neg p$: p is not on B, else B would be closed, so $v_B(p) = \mathrm{F}$ and $\neg p$ is true.
- Any other χ is used on B, since B is finished. So B passes through a node carrying the formulas of one of χ's ways. Each of them is shorter than χ: by the count of symbols in Section 4.3, $\neg\varphi$ has one symbol more than φ, and $(\varphi \circ \psi)$ has the symbols of φ and ψ plus 3, so each of φ, ψ, $\neg\varphi$, $\neg\psi$ is shorter than $(\varphi \circ \psi)$ and than its negation, and φ is shorter than $\neg\neg\varphi$. So each is true under v_B by the hypothesis, and then χ is true by the clause of its row in the bullets of (b), read from right to left.

The formulas of Δ are at the root, which is on B, so all of them are true under v_B.

The last sentence of the statement: apply (a), and then (b) or (c), to the list $\Delta = \Gamma, \neg\psi$. By Section 4.5, $\Gamma \models \psi$ exactly when this list is unsatisfiable, and a valuation making it true makes every formula of Γ true and ψ false.

■

Warning. A split needs **both** sides. Had we said, in the first tree, "suppose q is true", closed that branch and stopped, we would have shown only that no refuting row makes q true; the rows with q false and r true were never examined. A clash reached under a guess refutes the guess, not the claim. The converse slip: an open branch proves something only when it is **finished**. While a formula on it is unused, it may still close.

4.6.4 Choosing the next move

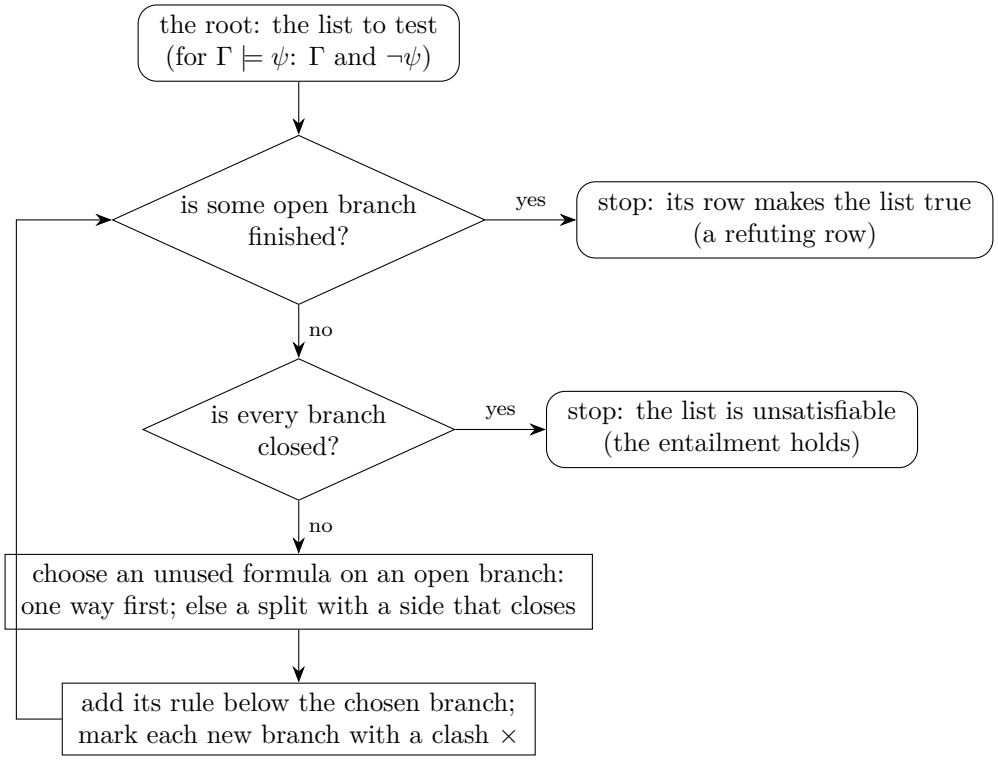
Every order gives the same verdict; the work differs. Five tactics keep a tree small.

1. **The conclusion first.** $\neg\psi$ is often one way, and it fixes the values that make the premises bite.
2. **One way before two ways.** A one-way rule adds to every branch at once; a split doubles

the work below it.

3. **Close early.** Before a split, look for a formula whose negation is already on the branch; after it, close every branch you can at once.
4. **Split where a side closes at once.** Among formulas that must split, prefer one with a side that clashes immediately.
5. **Stop at the first finished open branch.** One refuting row is enough.

The method as a chart:



The order of the tests decides nothing about the verdict (Proposition 4.6.1), only about the work.

Quick check. Search Example (A tree that closes) again, applying the rule of $p \to (q \vee r)$ before that of $\neg(p \to s)$. What changes, and what does not?

Solution. The one-way node p, $\neg s$ must now be written under each open side of the first split, at a stage of its own when that side is chosen, so the tree is larger. The side $\neg p$ closes only after p has been added under it. Every branch still closes, and by Proposition 4.6.1 the verdict cannot change.

4.6.5 Testing a list for satisfiability

The same tree works for any finite list. For a list of claims about an integer, an open finished branch gives a row, and often a number.

Example. Three claims about an integer

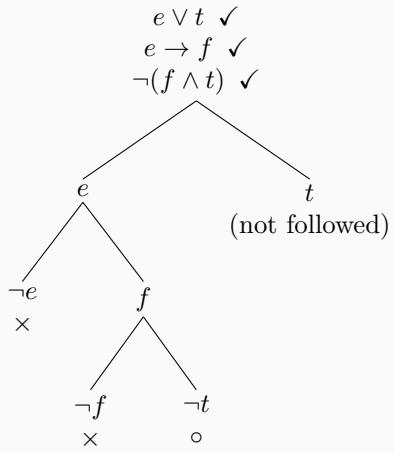
Let n be an integer. Can these three claims all be true?

(1) n is even or n is a multiple of 3.

(2) If n is even, then n is a multiple of 4.

(3) n is not both a multiple of 4 and a multiple of 3.

Solution. Take the atoms e: "n is even", t: "n is a multiple of 3" and f: "n is a multiple of 4". The list is $e \vee t$, $e \to f$, $\neg(f \wedge t)$. No rule is one way, so we split $e \vee t$: e or t. Under e, the rule of $e \to f$ gives $\neg e$, which closes, or f. Under f, the rule of $\neg(f \wedge t)$ gives $\neg f$, which closes, or $\neg t$. That branch, $e, f, \neg t$ with every compound formula used, is open and finished.



A number with its row (e, f true, t false) is $n = 4$: even, a multiple of 4, not a multiple of 3. All three claims hold for it. The branch t need not be followed.

The search treats e, f and t as independent letters. Here both rows that satisfy the list, $e, f, \neg t$ and $\neg e, \neg f, t$ (the branch t, followed to the end), happen to be rows of integers, 4 and 3; nothing in the search guarantees it. A row such as $e = \mathrm{F}$, $f = \mathrm{T}$, which no integer has (a multiple of 4 is even), is a row like any other to the search, as the warning at the end of Section 4.4 says.

Section 3.2 asked whether a list is **inconsistent**; the search asks whether it is **unsatisfiable**, and Sections 4.7 and 4.8 prove that the two questions always have the same answer.

Quick check. Section 3.2 derived $\bot$ from the list $p \to q$, $q \to \neg p$, p (Exercise (B4: Derive $\bot$, or make every member true, used later) (a)). Is it satisfiable?

Solution. No. The rule of $p \to q$ gives $\neg p$, which closes against p, or q. Under q, the rule of $q \to \neg p$ gives $\neg q$, which closes, or $\neg p$, which closes against p. Every branch is closed, so by Proposition 4.6.1 (b) the list is unsatisfiable.

Second reading · *A tree is a table with rows merged*

Each branch of a search tree stands for the rows that agree with its atoms and negated atoms and make true every formula split above it: a branch carrying p and $\neg s$, with no split above it, stands for every row with p true and s false. A closed branch stands for no row at all. A split keeps the rows of its branch that make the split formula true and divides them into two groups, which may overlap; a row of the branch that makes the formula false lies on neither side and is struck out at the split itself. So a tree whose branches all close is a truth table whose rows have been grouped and struck out a group at a time. In the first tree of this section, the root's p and $\neg s$ leave four of the sixteen rows on p, q, r, s as candidates; the branch through q and its closed leaves strike out the two of those with q true, the branch through r the one left with q false and r true, and the row with q and r both false lies on neither side of the split of $p \to (q \vee r)$, which it makes false. Between them, the root's atoms, the splits and the five closed branches strike out all sixteen rows as candidates for a refuting row, and nobody had to write any of them down.

4.6.6 The withheld verdicts

Section 3.2's derive-or-break items withheld the verdict, and a claim that resisted both deriving and breaking by hand stayed open. The search settles each **semantically**: a closed tree shows the claim valid, an open finished branch gives a refuting row, always in finitely many steps.

Here are the four claims that Exercise B3 of Section 3.2 broke (Exercise (B3: Derive or break, used later), used later), read as entailments, each with the row its search returns. $\times$ marks a side that closes.

Claim	Refuting row	The search
(a) $\neg p \vee \neg q \models \neg(p \vee q)$	$p = \mathrm{F}$, $q = \mathrm{T}$	root $\neg p \vee \neg q$, $\neg\neg(p \vee q)$; add $p \vee q$; split $\neg p \vee \neg q$; under $\neg p$, $p \vee q$ gives p $\times$ or q: open
(d) $p \to q,\ \neg p \to (r \vee q),\ \neg r \models p$	$p = \mathrm{F}$, $q = \mathrm{T}$, $r = \mathrm{F}$	$p \to q$: $\neg p$ or q; under $\neg p$, $\neg p \to (r \vee q)$ gives $\neg\neg p$ $\times$ or $r \vee q$; then r $\times$ or q: open
(f) $(p \to q) \to r,\ \neg r \models q$	$p = \mathrm{T}$, $q = \mathrm{F}$, $r = \mathrm{F}$	$\neg(p \to q)$ or r $\times$; $\neg(p \to q)$ gives $p, \neg q$: open
(g) $p \vee q,\ p \models \neg q$	$p = q = \mathrm{T}$	$\neg\neg q$ gives q; $p \vee q$: p or q, both open

Each row is the row of Section 3.2's break, whose statements about x take these values: in (a) they are $x > 2$ and $x > 0$ at $x = 1$; in (d), $x > 5$, $x > 0$ and $x > 10$ at $x = 1$; in (f), $x > 0$, $x > 2$ and $x > 3$ at $x = 1$; in (g), $x > 1$ and $x > 2$ at $x = 3$. Section 3.2 found these rows by hand; the search finds them without a guess, and would have said "valid" had there been none.

Two questions remain, and they are the next two sections': does a refuting row show that no derivation exists, and does a valid claim always have a derivation? (Sections 4.7 and 4.8.)

Remark (Looking ahead). Section 4.7 proves that a refuting row rules out every derivation. With it, a search that stays open on a subgoal tells a derivation search to stop chasing that subgoal: the move *test the subgoal*.

4.6.7 Exercises

4.6.7.1 A. Check your understanding

Exercise. A1

(a) Which of these formulas split when used: $p \wedge q$; $\neg(p \wedge q)$; $\neg(p \to q)$; $p \leftrightarrow q$; $\neg\neg p$?

(b) An open finished branch carries $p \to q$ (used), $\neg q$, $\neg p$, r. Which row does it give?

(c) True or false: if some branch of a finished tree for $\Gamma, \neg\psi$ is closed, then $\Gamma \models \psi$.

(d) True or false: two finished trees for the same list, built in different orders, can give different verdicts.

(e) Why does the warning say that a clash reached under a guess refutes the guess, not the claim?

Solution.

(a) $\neg(p \wedge q)$ and $p \leftrightarrow q$.

(b) $p = \mathrm{F}$, $q = \mathrm{F}$, $r = \mathrm{T}$, and every other atom F.

(c) False: every branch must be closed. One closed branch only rules out the rows of that branch.

(d) False. By Proposition 4.6.1, a closed tree means that no row makes the list true, and an open finished branch means that one does; both cannot hold.

(e) The guess picked one side of a split. The rows on the other side were never examined, and a refuting row may lie there.

Exercise. A2: Drill

The parts run from writing one rule, through closing a branch and reading off a row, to short searches.

(a) What does $\neg(p \vee q)$ add?

(b) What does $p \to (q \wedge r)$ add?

(c) What does $\neg(p \leftrightarrow q)$ add?

(d) What does $\neg\neg(p \wedge q)$ add?

(e) Is the branch $p \vee q,\ \neg p,\ \neg q,\ q$ closed?

(f) Is the branch $p \to q,\ \neg q,\ p$ closed?

(g) An open finished branch carries $\neg(p \wedge q)$ (used), p, $\neg q$. Which row does it give?

(h) Search: does $p \vee q \models p$ hold?

(i) Search: does $p \wedge q \models p \vee r$ hold?

(j) Is the list $p \leftrightarrow q,\ q \leftrightarrow \neg p$ satisfiable?

(k) For the list $p \vee q,\ \neg(r \to s)$, which formula should be used first, and why?

(l) "Suppose q is true; then the branch closes; so the claim is valid." Is this correct?

Solution.

(a) $\neg p$, $\neg q$ (one way).

(b) $\neg p$, or $q \wedge r$ (two ways).

(c) $p, \neg q$, or $\neg p, q$.

(d) $p \wedge q$.

(e) Yes: it holds q and $\neg q$.

(f) Not yet. Using $p \to q$ gives $\neg p$, which closes against p, or q, which closes against $\neg q$; so the tree below it closes.

(g) $p = \mathrm{T}$, $q = \mathrm{F}$.

(h) No. Root $p \vee q,\ \neg p$; split: p closes, and q is open and finished. The row $p = \mathrm{F}$, $q = \mathrm{T}$ refutes it.

(i) Yes. Root $p \wedge q,\ \neg(p \vee r)$; one way: p, q; one way: $\neg p, \neg r$; closed, so valid.

(j) No. The first gives p, q or $\neg p, \neg q$. Under each, the second gives $q, \neg p$ or $\neg q, \neg\neg p$, and all four branches close. The last one, $\neg p, \neg q, \neg q, \neg\neg p$, closes at once: $\neg\neg p$ is the negation of $\neg p$.

(k) $\neg(r \to s)$: it is one way.

(l) No: that is a guess. The case with q false must also close.

4.6.7.2 B. Practice

Exercise. B1: Valid or refuted?

Decide each claim by a search tree; you may stop at the first finished open branch. For a refuted claim, read off a refuting row and check it.

(a) $p \to q,\ r \to s,\ p \vee r \models q \vee s$

(b) $p \to q,\ r \to s,\ q \vee s \models p \vee r$

(c) $(p \wedge q) \to r \models (p \to r) \vee (q \to r)$

(d) $p \leftrightarrow q,\ q \vee r \models p \vee \neg r$

Solution. The verdict comes first, then the tree as its stages; $\times$ marks a side that closes.

(a) **Valid.** $\neg(q \vee s)$ adds $\neg q, \neg s$. $p \vee r$: p or r. Under p: $p \to q$ gives $\neg p$ $\times$ or q $\times$. Under r: $r \to s$ gives $\neg r$ $\times$ or s $\times$. Every branch is closed.

(b) **Refuted.** $\neg(p \vee r)$ adds $\neg p, \neg r$. $q \vee s$: q or s. Under q: $p \to q$ gives $\neg p$ or q; take $\neg p$. Then $r \to s$ gives $\neg r$ or s; take $\neg r$. The branch through $q, \neg p, \neg r$ is open and finished, and the search stops there; the other sides are not followed. Row $p = \mathrm{F}$, $q = \mathrm{T}$, $r = \mathrm{F}$, $s = \mathrm{F}$: both implications have false hypotheses, $q \vee s$ is T, and $p \vee r$ is F.

(c) **Valid.** $\neg((p \to r) \vee (q \to r))$ adds $\neg(p \to r), \neg(q \to r)$; these add $p, \neg r$ and $q, \neg r$. $(p \wedge q) \to r$: $\neg(p \wedge q)$ or r $\times$. $\neg(p \wedge q)$: $\neg p$ $\times$ or $\neg q$ $\times$.

(d) **Refuted.** $\neg(p \vee \neg r)$ adds $\neg p, \neg\neg r$, and $\neg\neg r$ adds r. $p \leftrightarrow q$: p, q $\times$ (against $\neg p$) or

$\neg p, \neg q$. $q \vee r$: q $\times$ or r. The branch $\neg p, \neg q, r$ is open and finished. Row $p = \mathrm{F}$, $q = \mathrm{F}$, $r = \mathrm{T}$: $p \leftrightarrow q$ is T, $q \vee r$ is T, and $p \vee \neg r$ is F.

Claims (a) and (b) are look-alikes, with the premises' disjunction and the conclusion traded, and only (a) holds.

Exercise. B2: Can all five be true?

Let n be an integer, and consider:

(1) if n is divisible by 6, then n is divisible by 4;
(2) n is divisible by 6, or n is odd;
(3) n is not odd;
(4) if n is divisible by 4, then n is not divisible by 3;
(5) if n is divisible by 6, then n is divisible by 3.

(a) With the atoms s, f, o, t for "divisible by 6", "divisible by 4", "odd", "divisible by 3", decide by a search tree whether the list of five formulas is satisfiable. (b) Drop one claim so that the rest are satisfiable, and give an integer for which they all hold.

Solution.

(a) The list is $s \to f$, $s \vee o$, $\neg o$, $f \to \neg t$, $s \to t$. No rule is one way. Split $s \vee o$, which has a side that closes at once: s, or o $\times$ (against $\neg o$). Under s: $s \to f$ gives $\neg s$ $\times$ or f; $f \to \neg t$ gives $\neg f$ $\times$ or $\neg t$; $s \to t$ gives $\neg s$ $\times$ or t $\times$ (against $\neg t$). Every branch closes: the list is unsatisfiable, so no integer makes all five claims true, and their form alone forbids it.

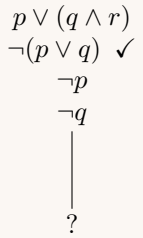
(b) Drop (4). In the tree without it, the branch s, f, t, with o false, is open and finished. The integer $n = 12$ is divisible by 6, 4 and 3, and not odd, so it satisfies (1), (2), (3) and (5). Dropping (1) instead also works: under s, the rule of $f \to \neg t$ gives $\neg f$, then $s \to t$ gives t, and the branch $s, \neg f, t$, with o false, is open and finished; $n = 6$ realizes it.

Marking scheme. (6 points)

- 1: the formulas.
- 3: the closed tree.
- 2: (b), with an integer checked against each remaining claim.

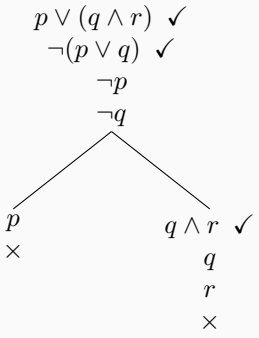
Exercise. B3: Name the move, then complete it

A search for $p \vee (q \wedge r) \models p \vee q$ has reached this stage.



Name the move, then complete it.

Solution. No one-way formula is left unused, so the move is to split the disjunction $p \vee (q \wedge r)$. The side p closes against $\neg p$. The side $q \wedge r$ is one way: it adds q, r, and closes against $\neg q$. Every branch is closed: the entailment holds.



Marking scheme. (4 points)

- 1: the move named (split, since nothing one-way is left).
- 2: the completed tree.
- 1: the verdict.

4.6.7.3 C. Going deeper

Exercise. C1: Rules for a new connective

Add the connective $\mid$ ("not both"), with the clause (V7) of Exercise (C1: Substitution with a new connective, the last half blank, used later): $\varphi \mid \psi$ is F exactly when both parts are T.

(a) Read search rules for $\varphi \mid \psi$ and for $\neg(\varphi \mid \psi)$ off (V7), and say which is one way.

(b) Decide $p \mid q,\ p \models \neg q$ by a search.

(c) Decide by a search whether $p \mid (p \mid p)$ is a tautology.

Solution.

(a) $\varphi \mid \psi$ is true unless both parts are true, so it is true in two ways: it adds $\neg\varphi$ or $\neg\psi$. $\neg(\varphi \mid \psi)$ is true exactly when both parts are true, one way: it adds φ, ψ.

(b) Root $p \mid q,\ p,\ \neg\neg q$. The formula $\neg\neg q$ adds q. Then $p \mid q$ gives $\neg p$ × or $\neg q$ ×. Every branch is closed: valid.

(c) The root is the negated formula, $\neg(p \mid (p \mid p))$. It is one way: it adds p and $p \mid p$. The latter gives $\neg p$ × or $\neg p$ ×. The tree is closed, so the negation is unsatisfiable and the

formula is a tautology. It says "not both p and not p", since $p \mid p$ is true exactly when p is false.

Marking scheme. (8 points)

- 3: (a), with each rule tied to its rows of (V7).
- 2: (b).
- 3: (c), including the root as the negated formula.

Exercise. C2: Exit check: refuting-row search

Close the section first. For each claim, run a search tree, state the tactics you used in your order of moves, give the verdict, and say which part of Proposition 4.6.1 justifies it.

(a) $p \to (q \vee r),\ \neg r \models p \to q$

(b) $p \vee q,\ p \to r \models q \vee \neg r$

Solution.

(a) The conclusion first: $\neg(p \to q)$ adds $p, \neg q$. Then $p \to (q \vee r)$, which has a side that closes at once: $\neg p$ $\times$ or $q \vee r$. Then $q \vee r$: q $\times$, or r $\times$ (against $\neg r$). Every branch is closed, so by (b) the counterexample set is unsatisfiable, and the entailment holds.

(b) The conclusion first: $\neg(q \vee \neg r)$ adds $\neg q, \neg\neg r$, and then $\neg\neg r$ adds r (one way before two ways). Split $p \vee q$, whose side q closes at once: p, or q $\times$. Under p: $p \to r$ gives $\neg p$ $\times$ or r, and the branch is open and finished (stop at the first finished open branch). By (c), its row $p = \mathrm{T}$, $q = \mathrm{F}$, $r = \mathrm{T}$ makes every root formula true: $p \vee q$ is T, $p \to r$ is T, and $q \vee \neg r$ is F. The claim is refuted.

Marking scheme. (10 points)

- 1 + 1: the roots, which are the counterexample sets.
- 3 + 3: the trees.
- 1: the tactics named.
- 1: the parts of Proposition 4.6.1 cited correctly, (b) for (a) and (c) for (b).

Fatal errors (0 for the part): a root without the negated conclusion; a verdict from an unfinished open branch.

Exercise. C3: Challenge: the search with one more connective

Extend search trees by the rules of C1 for $\mid$. Prove that Proposition 4.6.1 still holds for lists of formulas of the extended language: give the new cases of each part of its proof.

Solution.

(a) The new rules add immediate parts or their negations, so every added formula is still a subformula of a root formula or the negation of one, and the count of the proof is unchanged.

(b) New cases of the claim. If v makes $\varphi \mid \psi$ true, then by (V7) not both parts are true,

so v makes $\neg\varphi$ or $\neg\psi$ true: one of the two new nodes. If v makes $\neg(\varphi \mid \psi)$ true, then $\varphi \mid \psi$ is false, so both parts are true: the one new node.

(c) New cases of the induction. A used $\varphi \mid \psi$ on B has $\neg\varphi$ or $\neg\psi$ on B, which is shorter and true by the hypothesis, so one part of $\varphi \mid \psi$ is false and $\varphi \mid \psi$ is true by (V7). A used $\neg(\varphi \mid \psi)$ has φ and ψ on B, both shorter and true by the hypothesis, so $\varphi \mid \psi$ is false and its negation true. The base cases are unchanged: an atom, a negated atom, $\bot$ and $\neg\bot$.

Marking scheme. (10 points)

- 2: (a).
- 4: (b), both rules.
- 4: (c), both rules, with the hypothesis applied to shorter formulas.

Fatal error: in (c), using the hypothesis for $\varphi \mid \psi$ itself.

4.7 Soundness

Chapter 3 ended on a question it could not answer. Section 3.1 searched for a derivation of $P \to Q \vdash Q \to P$, and every route came back to the same gap; but a search that fails proves nothing, since a cleverer derivation might exist. The row that refutes affirming the consequent in Section 4.5, $p = \mathrm{F}$ and $q = \mathrm{T}$, finds what the search could not: it makes $p \to q$ true and $q \to p$ false, and the search of Section 4.6 finds such rows whenever they exist. This section proves the theorem that turns such a row into a proof that **no** derivation exists. It is called soundness: whatever the rules derive is true in every row where the premises are true. The proof is the book's first long proof about proofs, and it is an induction on the lines of a derivation.

4.7.1 What must be proved, and why it is not obvious

Remark. **Two levels.** This section speaks about the calculus from outside it.

Object level: written in the language	Meta level: claims about it, proved in ordinary mathematics
formulas; derivations, with their lines, boxes, justifications and printed purposes ("asm, for →I", which the proof reads as the rule whose case applies)	"$\Gamma \vdash \psi$", "$\Gamma \models \psi$", "line k is justified by ∨E", "$\Gamma_k \models \varphi_k$"

Section 4.2's checklist reads: objects, the lines of one derivation; property, $\Gamma_k \models \varphi_k$ (contexts are defined below); base cases, the premises and assumptions, which cite nothing; steps, one for each rule; principle, Theorem 4.2.2.

Second reading · *Is the proof circular?*

The proof below argues by cases ("if $v(\varphi) = \mathrm{F}$ ... if $v(\varphi) = \mathrm{T}$") and by contradiction, in the metalanguage: the very patterns that $\vee$E and RAA record in the object language. That is not circular. The theorem is not meant to persuade someone who rejects reasoning by cases. It is a theorem of ordinary mathematics about a calculus, and what it rules out is a mismatch between the rules and the tables, of the kind the connective in the fold of Section 4.4 has.

Chapters 1 and 2 checked each rule as it arrived. $\wedge$I is safe because two true statements make a true conjunction (Section 1.2); $\to$E is safe because that is what "if ... then" asserts (Section 1.3). Section 1.2 said once that Chapter 4 would prove it for every rule of Chapters 1 and 2 at once. Why is that more than collecting the separate checks?

The first attempt at a proof is: go down a derivation, line by line, and show that each line is true in every row in which the premises are true. For a derivation without boxes this works. Each line is a premise, or comes from earlier lines by a rule, and the rule turns true lines into a true line. But it fails at the first box. Here is the derivation of Proposition 2.3.2, written with atoms:

1	$p \to q$	
2	$\quad \neg q$	asm, for $\to$I
3	$\quad\quad p$	asm, for $\neg$I
4	$\quad\quad q$	$\to$E 1, 3
5	$\quad\quad \bot$	$\neg$E 4, 2
6	$\quad \neg p$	$\neg$I 3–5
7	$\neg q \to \neg p$	$\to$I 2–6

Line 3 is p, and the row $p = \mathrm{F}$, $q = \mathrm{F}$ makes the premise $p \to q$ true and line 3 false. Line 5 is $\bot$, which is false in **every** row. So the lines inside a box are not true in the rows where the premises are true, and the first attempt collapses. Yet the derivation is correct, and its last line $\neg q \to \neg p$ is true in every row making $p \to q$ true (Example (An implication and its contrapositive)).

What goes wrong is that a line inside a box is not claimed on the strength of the premises alone. Line 4 says: *if* $p \to q$, $\neg q$ and p all hold, then q holds. That is a true claim about rows. Every line depends on the premises **and** on the assumptions of the boxes around it, and a correct statement about each line must mention both. Below we give this list a name: the context of the line.

4.7.2 Derivations of formulas

From now on the lines of a derivation carry **formulas** in the sense of Definition 4.1.1, with atoms $p, q, r, \ldots$. Nothing in the rules changes. Every rule was stated as a schema, whose letters φ, ψ, χ stand for arbitrary statements (Definition 1.2.1), so it applies to formulas as it stands, and Definition 1.4.2 and Definition 1.2.3 read the same with "formula" for "statement". As Section 4.1 said, a derivation written with Chapter 2's letters P, Q is a pattern, and reading its letters as atoms, or as any formulas, gives a derivation of formulas.

Two conventions for the rest of the section. The **primitive rules** are those of Chapters 1 and 2 collected in Section 3.1: R, $\wedge$I, $\wedge$E, $\to$I, $\to$E, $\vee$I, $\vee$E, $\neg$I, $\neg$E, $\bot$E, RAA, $\leftrightarrow$I and $\leftrightarrow$E. The **box rules** among them are the five that cite a box: $\to$I, $\neg$I, RAA, $\vee$E and $\leftrightarrow$I. The other eight, R included, cite only lines; we call them the **one-step rules**.

A derivation may also use derived rules such as MT and LEM. A line citing one abbreviates a stretch of derivation in rules proved earlier, with the letters replaced by formulas (Definition 2.3.1), and Section 2.3 showed how to write it out: repeat the cited lines, then paste the derivation of the derived rule below them. Every chain of derived rules ends in primitive rules. So whenever $\Gamma \vdash \psi$, there is also a derivation of ψ from premises in Γ that uses **primitive rules only**. That is the only kind of derivation the proof below has to consider.

4.7.3 The context of a line

Let a derivation have lines $1, \dots, n$. We write φ_k for the formula on line k. The **length** of a derivation is its number of lines. The **context** of line k is the list of the premises together with the assumptions of every box containing line k: the statements the line depends on. We write it Γ_k. The context of the last line is just the premises, since by Definition 1.4.2 the last line lies in no box. For lists of formulas Δ and Δ', we write $\Delta \subseteq \Delta'$ when every formula of Δ occurs in Δ'.

Exercise C4 of Section 1.4 (Exercise (C4: The assumptions a line depends on, used later)) wrote such lists by hand, as sets of line numbers; its part (c) observed what Lemma 4.7.1 below proves.

The context of a line is everything that line depends on: the premises, and every assumption still open where the line stands.

In the derivation of Proposition 2.3.2 above, the boxes are 2–6 and 3–5, and the contexts are:

Line	Formula	Boxes containing it	Context
1	$p \to q$	none	$p \to q$
2	$\neg q$	2–6	$p \to q,\ \neg q$
3, 4, 5	$p,\ q,\ \bot$	2–6, 3–5	$p \to q,\ \neg q,\ p$
6	$\neg p$	2–6	$p \to q,\ \neg q$
7	$\neg q \to \neg p$	none	$p \to q$

Quick check. In that derivation, which lines have p in their context?

Solution. Lines 3, 4 and 5, the lines of the box opened by p.

Now look at each line through its context. Line 3 is one of the formulas in its own context, so every row making the context true makes line 3 true. Line 4: every row making $p \to q$ and p true makes q true. Line 5: no row makes $p \to q$, $\neg q$, and p all true (the first and third force q), so the context is unsatisfiable and it entails even $\bot$ (Section 4.5). Line 6: every row making $p \to q$ and $\neg q$ true makes p false. Line 7: its context is the premise, and we know $p \to q \models \neg q \to \neg p$. So on this derivation, every line k satisfies

$$\Gamma_k \models \varphi_k.$$

This is the statement that survives the boxes, and it is the one we shall prove for every line of every derivation.

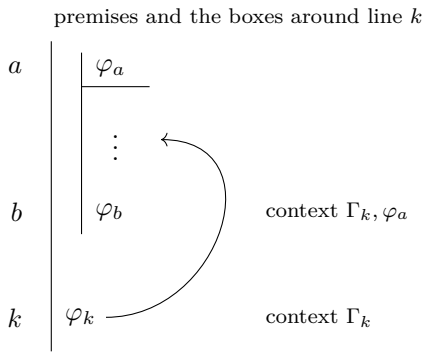
Two facts about contexts, read off the scope rule, do the bookkeeping. The first says that citing an available line never involves an assumption the citing line does not have. The second says exactly what the context of a box's last line is, when the box is cited by a box rule.

Lemma 4.7.1. Contexts in a Derivation

Let lines i, k and $a \leq b$ belong to a derivation, and let Γ_j denote the context of line j.

(a) If line i is available at line k, then $\Gamma_i \subseteq \Gamma_k$.

(b) If line k is justified by a box rule citing the box whose first line is a and whose last line is b, then the formulas of Γ_b are exactly those of Γ_k together with φ_a, the assumption of that box. We write this $\Gamma_b = \Gamma_k, \varphi_a$.



Part (b): the last line of a cited box depends on everything line k depends on, and on the box's own assumption. The arrow is the citation.

Idea. Both parts compare two sets of boxes, since contexts differ only in the assumptions of the boxes around a line. For (a), the scope rule says that the boxes around line i are among the boxes around line k. For (b), the boxes around line b should be the cited box itself and the boxes around line k; the danger is a further box inside the cited box that also ends on line b, and the box rules forbid exactly that.

The proof reads both parts off the scope rule; it is in the fold below.

Second reading · *The proof: reading the scope rule*

Proof. (a) The premises belong to every context. Any other formula of Γ_i is the assumption of a box containing line i. Since line i is available at line k, every box containing line i also contains line k, by the scope rule of Definition 1.4.2. So that assumption belongs to Γ_k.

(b) Let B be the box from line a to line b. Since line k cites B, the box B is available at line k. By the scope rule, it ends before line k, and the boxes containing B are exactly the boxes containing line k.

We find the boxes containing line b. Every box containing line k contains B, and so contains

line b; and B contains line b. Conversely, let C be a box containing line b. By Definition 1.4.2, two boxes are either disjoint or one inside the other; C and B share line b, so $C = B$, or C contains B, or C lies inside B. The last is impossible: then line b, the last line of B, would be a line of a box inside B, and every box rule requires the last line of its box to be a line of that box itself, not of a box inside it (Definition 1.4.1, Definition 2.2.1, Definition 2.5.1, Definition 1.6.1, Definition 2.4.1). So $C = B$, or C is one of the boxes containing B, which are the boxes containing line k.

Hence the boxes containing line b are B and the boxes containing line k. The assumption of B is φ_a, so Γ_b consists of the premises, the assumptions of the boxes containing line k, and φ_a, which is Γ_k together with φ_a.

■

In the derivation of Proposition 2.3.2, line 6 cites the box 3–5 by ¬I, and $\Gamma_5 = p \to q,\ \neg q,\ p$ is $\Gamma_6 = p \to q,\ \neg q$ with the assumption p added, as (b) says. Line 5 cites lines 4 and 2, both available there, and Γ_4 and Γ_2 are contained in Γ_5, as (a) says.

↻ **Second reading** · *A box that ends inside another*

Warning. Part (b) needs the box to be **cited by a box rule**. Definition 1.4.2 allows a box to end on the same line as a box inside it, and then its last line has a larger context. In the derivation below, the box 2–4 is available at line 5, but $\Gamma_4 = p,\ q,\ r$ contains r, the assumption of the inner box, as well as q. No box rule may cite 2–4: its last line lies inside the box 3–4, not in 2–4 itself. (This is why the derivation of Example (Regrouping a disjunction) needs its line 11.)

1	p	
2	$\quad q$	
3	$\quad\quad r$	
4	$\quad\quad p$	R 1
5	p	R 1

Quick check. In the derivation of Proposition 2.3.2 above, which case of the soundness claim handles line 5? Which lines' contexts are involved, and why does $\Gamma_5 \models \bot$ hold?

Solution. Line 5 is ⊥, by ¬E citing lines 4 and 2: a one-step rule. Lines 4 and 2 are available at line 5, so by Lemma 4.7.1 (a), $\Gamma_4 = p \to q,\ \neg q,\ p$ and $\Gamma_2 = p \to q,\ \neg q$ are contained in $\Gamma_5 = p \to q,\ \neg q,\ p$. The claim for the earlier lines gives $\Gamma_4 \models q$ and $\Gamma_2 \models \neg q$. So a valuation making Γ_5 true would make both q and $\neg q$ true, which clause (V2) of Definition 4.4.2 forbids.

No valuation makes Γ_5 true, and $\Gamma_5 \models \bot$ holds because there is no row to check.

4.7.4 The soundness theorem

We can now state the theorem that Chapters 1 and 2 kept pointing to.

Theorem 4.7.2. Soundness

Let Γ be a list of formulas and ψ a formula. If $\Gamma \vdash \psi$, then $\Gamma \models \psi$.

Try it first (ten minutes): suppose line k is $\varphi \to \psi$, by $\to$I from a box that opens with φ and ends with ψ. Suppose you already know that every row making the context of line k and φ true makes ψ true. Why is $\varphi \to \psi$ true in every row making the context of line k true? What happens in the rows where φ is false?

Idea. An attempt at the $\to$I case runs into two things. First, knowledge about the box's last line is knowledge about rows making its **own** context true, which contains the assumption φ; that is why the claim must carry contexts, and why Lemma 4.7.1 (b) is needed. Second, the rows in which φ is false are not covered by the box at all, and they need nothing: there $\varphi \to \psi$ is true by (V5). The same two moves handle every box rule. The Step roadmap:

① Prove the stronger claim $\Gamma_k \models \varphi_k$ for **every** line k of a derivation in primitive rules.

② Use strong induction on k (Theorem 4.2.2): a rule may cite **any** earlier line, not only the line just above, so the hypothesis must be available for all earlier lines at once.

③ Split on how line k is justified. A premise or an assumption belongs to its own context. A one-step rule: apply the hypothesis to the cited lines, move their contexts into Γ_k by Lemma 4.7.1 (a), and use one clause of Definition 4.4.2. A box rule: apply the hypothesis to each box's last line, whose context is Γ_k plus the box's assumption by Lemma 4.7.1 (b), and argue about the rows where that assumption is true or false.

④ The last line lies in no box, so its context is the premises, all in Γ; adding the other formulas of Γ keeps the entailment.

Before reading on: in the case of RAA *(Claim 3 below), where does the proof use that $v(\neg\varphi)$ is* F *exactly when $v(\varphi)$ is* T*, and which half of that fact does it use?*

Proof. **Step 1 (reduction to primitive rules).** Suppose $\Gamma \vdash \psi$. As shown in the subsection on derivations of formulas, there is a derivation of ψ whose premises all belong to Γ and which uses primitive rules only. Fix one, with lines $1, \ldots, n$; let φ_k be the formula on line k and Γ_k its context.

Step 2 (the induction). For $k \in \mathbb{N}$, let $S(k)$ be the statement: *if* $1 \le k \le n$, *then* $\Gamma_k \models \varphi_k$. We prove $S(k)$ for every k by Theorem 4.2.2. So let $k \in \mathbb{N}$, and assume $S(j)$ for every $j < k$. If $k = 0$ or $k > n$, there is nothing to prove, so let $1 \le k \le n$. Every line and every box that line k cites is available at line k, and so lies before line k (Definition 1.4.2). Hence the induction hypothesis $\Gamma_j \models \varphi_j$ holds for every line j that line k cites, and for the last line of every box it cites.

Let v be a valuation that makes every formula of Γ_k true. We show $v(\varphi_k) = \mathrm{T}$; since v is arbitrary, this proves $\Gamma_k \models \varphi_k$. By Definition 1.4.2, line k is a premise, an assumption, or the conclusion of one primitive rule, and the three claims below treat these cases.

Step 3 (the cases).

Claim. If line k is a premise or an assumption, then $v(\varphi_k) = \mathrm{T}$.

Proof of Claim. A premise belongs to every context. An assumption opens a box, that box contains line k, and so its assumption φ_k belongs to Γ_k. In both cases φ_k is one of the formulas that v makes true.

■

For the one-step rules we first record what the hypothesis gives for a cited line. If line k cites line i, then line i is available at line k, so $\Gamma_i \subseteq \Gamma_k$ by Lemma 4.7.1 (a). Hence v makes every formula of Γ_i true, and the induction hypothesis $\Gamma_i \models \varphi_i$ gives

$$v(\varphi_i) = \mathrm{T} \quad \text{for every line } i \text{ cited by line } k. \tag{$*$}$$

Claim. If line k is justified by a one-step rule, then $v(\varphi_k) = \mathrm{T}$.

Proof of Claim. We take the rules one at a time; in each, the cited lines are true under v by $(*)$.

R. Line k repeats a line i, so $\varphi_k = \varphi_i$, and $v(\varphi_k) = \mathrm{T}$.

$\wedge$I. Line k is $\varphi \wedge \psi$, citing lines holding φ and ψ. Both are true under v, so $\varphi \wedge \psi$ is true by (V3).

$\wedge$E. Line k is φ, citing a line holding $\varphi \wedge \psi$. That line is true under v, so by (V3) both conjuncts are true, and in particular φ. The second form, which writes ψ, is the same with the other conjunct, and we skip it.

$\to$E. Line k is ψ, citing lines holding $\varphi \to \psi$ and φ. Both are true under v. By (V5), an implication with a true hypothesis and a false conclusion is false; since $\varphi \to \psi$ is true and φ is true, ψ is true.

$\vee$I. Line k is $\varphi \vee \psi$ or $\psi \vee \varphi$, citing a line holding φ. That line is true under v, so both disjunctions are true by (V4).

$\neg$E. Line k is $\bot$, citing lines holding φ and $\neg\varphi$. By $(*)$ both would be true under v, and by (V2) $\neg\varphi$ is false whenever φ is true. So no valuation makes every formula of Γ_k true, and there is no v to consider: $\Gamma_k \models \bot$ holds because there is no row to check.

$\bot$E. Line k is any formula φ, citing a line holding $\bot$. By $(*)$, $v(\bot) = \mathrm{T}$, which (V1) forbids. So again no valuation makes every formula of Γ_k true, and $\Gamma_k \models \varphi$ holds.

$\leftrightarrow$E. Line k is ψ, citing lines holding $\varphi \leftrightarrow \psi$ and φ. Both are true under v; by (V6), $v(\psi) = v(\varphi) = \mathrm{T}$. The second form, which writes φ from $\varphi \leftrightarrow \psi$ and ψ, is the same with the sides exchanged, and we skip it.

■

For the box rules we record what the hypothesis gives for a cited box. If line k cites, by a box rule, the box from line a to line b, then $\Gamma_b = \Gamma_k, \varphi_a$ by Lemma 4.7.1 (b), and the induction

hypothesis for line b says

$$\Gamma_k, \varphi_a \models \varphi_b : \quad \text{every valuation making } \Gamma_k \text{ and } \varphi_a \text{ true makes } \varphi_b \text{ true.} \qquad (**)$$

Claim. If line k is justified by a box rule, then $v(\varphi_k) = \mathrm{T}$.

Proof of Claim. Recall that v makes every formula of Γ_k true.

$\to$I. Line k is $\varphi \to \psi$, citing a box that opens with φ and ends with ψ. By $(**)$, every valuation making Γ_k and φ true makes ψ true. If $v(\varphi) = \mathrm{F}$, then $v(\varphi \to \psi) = \mathrm{T}$ by (V5). If $v(\varphi) = \mathrm{T}$, then v makes Γ_k and φ true, so $v(\psi) = \mathrm{T}$, and again $v(\varphi \to \psi) = \mathrm{T}$ by (V5).

$\neg$I. Line k is $\neg\varphi$, citing a box that opens with φ and ends with $\bot$. By $(**)$, every valuation making Γ_k and φ true makes $\bot$ true; since (V1) makes $\bot$ false under every valuation, no valuation makes Γ_k and φ true. As v makes Γ_k true, it must make φ false, and then $v(\neg\varphi) = \mathrm{T}$ by (V2).

RAA. Line k is φ, citing a box that opens with $\neg\varphi$ and ends with $\bot$. By $(**)$ and (V1), as in the previous case, no valuation makes Γ_k and $\neg\varphi$ true. Suppose $v(\varphi) = \mathrm{F}$. Then $v(\neg\varphi) = \mathrm{T}$ by (V2), so v makes Γ_k and $\neg\varphi$ true, which is impossible. Hence $v(\varphi) = \mathrm{T}$.

$\vee$E. Line k is χ, citing a line i holding $\varphi \vee \psi$, a box that opens with φ and ends with χ, and a box that opens with ψ and ends with χ. By $(*)$, $v(\varphi \vee \psi) = \mathrm{T}$, so by (V4), $v(\varphi) = \mathrm{T}$ or $v(\psi) = \mathrm{T}$. By $(**)$ for the first box, every valuation making Γ_k and φ true makes χ true; for the second, every valuation making Γ_k and ψ true makes χ true. If $v(\varphi) = \mathrm{T}$, then v makes Γ_k and φ true, so $v(\chi) = \mathrm{T}$. If $v(\psi) = \mathrm{T}$, the second box gives $v(\chi) = \mathrm{T}$ in the same way.

$\leftrightarrow$I. Line k is $\varphi \leftrightarrow \psi$, citing a box that opens with φ and ends with ψ, and one that opens with ψ and ends with φ. By $(**)$, every valuation making Γ_k and φ true makes ψ true, and every valuation making Γ_k and ψ true makes φ true. If $v(\varphi) = \mathrm{T}$, the first box gives $v(\psi) = \mathrm{T}$. If $v(\varphi) = \mathrm{F}$, then $v(\psi) = \mathrm{F}$ too, since $v(\psi) = \mathrm{T}$ would give $v(\varphi) = \mathrm{T}$ by the second box. In both cases $v(\varphi) = v(\psi)$, so $v(\varphi \leftrightarrow \psi) = \mathrm{T}$ by (V6).

■

The three claims cover every way line k can be justified, so $S(k)$ holds. By Theorem 4.2.2, $S(k)$ holds for every $k \in \mathbb{N}$: every line k of the derivation satisfies $\Gamma_k \models \varphi_k$.

Step 4 (the last line). The last line n of a derivation lies in no box (Definition 1.4.2), so its context Γ_n consists of the premises alone, and $\varphi_n = \psi$. By the induction of Steps 2 and 3, $\Gamma_n \models \psi$. Every premise belongs to Γ (Definition 1.2.3), so a valuation making every formula of Γ true makes every formula of Γ_n true, and hence makes ψ true. This proves $\Gamma \models \psi$.

■

The proof is long, but each case is one step of the truth-table reasoning of Section 4.4, placed in the right context. Every primitive rule has passed. So has every derived rule, since each abbreviates primitive ones. *The rules never lead from truth to falsehood*, not merely line by line but for every derivation that can ever be written.

↻ **Second reading** · *Why induction on the position of a line*

Why strong induction and the position k of a line? A rule may cite any available line, however

far back: line 7 of the derivation of Proposition 2.3.2 cites the box that begins at line 2. So the hypothesis is needed for all earlier lines at once. A different route would induct on the length of the whole derivation, removing its last line. But the lines above the last line need not form a derivation, since a box may still be open there (in the derivation of Proposition 2.3.2, lines 1–6 end inside the box 2–6), and that route has to cut boxes open. Carrying the context of every line avoids the surgery.

Example. Running the claim on a case split

Here is a derivation of $p \vee q \vdash q \vee p$. List the context of each line, and check $\Gamma_k \models \varphi_k$ for each k, naming the case of the proof that handles it.

1	$p \vee q$	
2	$\quad p$	asm, for $\vee$E
3	$\quad q \vee p$	$\vee$I 2
4	$\quad q$	asm, for $\vee$E
5	$\quad q \vee p$	$\vee$I 4
6	$q \vee p$	$\vee$E 1, 2–3, 4–5

Solution. The boxes are 2–3 and 4–5.

Line	Context Γ_k	Case	Why $\Gamma_k \models \varphi_k$
1	$p \vee q$	premise	φ_1 is in its context
2	$p \vee q,\ p$	assumption	φ_2 is in its context
3	$p \vee q,\ p$	$\vee$I citing 2	$\Gamma_2 \subseteq \Gamma_3$; a row making p true makes $q \vee p$ true
4	$p \vee q,\ q$	assumption	φ_4 is in its context
5	$p \vee q,\ q$	$\vee$I citing 4	as for line 3, with q
6	$p \vee q$	$\vee$E citing 1 and both boxes	see below

For line 6: by Lemma 4.7.1 (b), $\Gamma_3 = \Gamma_6, p$ and $\Gamma_5 = \Gamma_6, q$, and the claim for lines 3 and 5 says that rows making $p \vee q$ and p true, and rows making $p \vee q$ and q true, make $q \vee p$ true. A row making Γ_6 true makes $p \vee q$ true, so it makes p or q true, and one of the two boxes applies. Line 6 is the last line, and its context is the premise: so $p \vee q \models q \vee p$, as soundness promises.

4.7.5 Every rule, checked against the tables

Rule	Introduced	Case	Clause	Wrong version; its derivation; refuting row
R	1.4 (clause (D3))	Claim 2	none: the line is repeated	
$\wedge$I, $\wedge$E	1.2	Claim 2	(V3)	
$\rightarrow$E	1.3	Claim 2	(V5)	
$\rightarrow$I	1.4	Claim 3	(V5)	"a box from φ to ψ gives $\psi \rightarrow \varphi$"; the box $p \wedge q$, p would give $\vdash p \rightarrow (p \wedge q)$; row $p = \mathrm{T}$, $q = \mathrm{F}$
$\vee$I	1.5	Claim 2	(V4)	
$\vee$E	1.6	Claim 3	(V4)	"the two boxes may end in different lines; write the first box's"; from $p \vee q$, the boxes p and q would give $p \vee q \vdash p$; row $p = \mathrm{F}$, $q = \mathrm{T}$
$\neg$E, $\bot$E	2.1	Claim 2	(V2); (V1)	
$\neg$I	2.2	Claim 3	(V1), (V2)	"a box from φ to any line gives $\neg\varphi$"; the one-line box p would give $\vdash \neg p$; row $p = \mathrm{T}$
$\leftrightarrow$I, $\leftrightarrow$E	2.4	Claim 3; Claim 2	(V6)	
RAA	2.5	Claim 3	(V1), (V2)	"a box from φ to $\bot$ gives φ"; from $\neg p$, the box p, $\bot$ (by $\neg$E) would give $\neg p \vdash p$; row $p = \mathrm{F}$

The thirteen primitive rules, with the case and clause each uses. Each wrong version changes one condition, and its case fails there: the row makes the premises true and the line false.

↻ **Second reading** · *The four wrong derivations, written out*

Here are the four derivations the wrong versions would license, written as patterns with Chapter 2's letters; read P and Q as the atoms p and q for the rows of the table. Each is correct except at its last line, which only the wrong version allows. Wrong $\rightarrow$I reverses the implication its box proves; wrong $\vee$E closes on a line that only the first box reaches; wrong $\neg$I closes a box that

never reached $\bot$; wrong RAA opens its box with φ instead of $\neg\varphi$.

1	$P \wedge Q$	asm, for $\to$I
2	P	$\wedge$E 1
3	$P \to (P \wedge Q)$	$\to$I 1–2

1	$P \vee Q$	
2	P	asm, for $\vee$E
3	Q	asm, for $\vee$E
4	P	$\vee$E 1, 2–2, 3–3

1	P	asm, for $\neg$I
2	$\neg P$	$\neg$I 1–1

1	$\neg P$	
2	P	asm, for RAA
3	$\bot$	$\neg$E 2, 1
4	P	RAA 2–3

4.7.6 Refuting with one row

Soundness is most often used backwards. If the conclusion can be false while the premises are true, no derivation can exist.

Corollary 4.7.3. Refuting Row

Let Γ be a list of formulas and ψ a formula. If some valuation makes every formula of Γ true and ψ false, then $\Gamma \nvdash \psi$.

Proof. Let v make every formula of Γ true and ψ false. Then $\Gamma \models \psi$ fails (Definition 4.5.3). If $\Gamma \vdash \psi$ held, Theorem 4.7.2 would give $\Gamma \models \psi$. So $\Gamma \vdash \psi$ does not hold. ■

Here $\Gamma \nvdash \psi$ means that there is **no** derivation of ψ from premises in Γ: not with the primitive rules, not with any derived rule, not in a thousand lines. One row has settled a question about infinitely many derivations.

Quick check. Does the row $p = q =$ T refute $p \to q \vdash q$? Does the row $p = q =$ F?

Solution. The first does not, since its conclusion q is true. The second does: $p \to q$ is T and q is F. So $p \to q \nvdash q$, by Corollary 4.7.3.

Chapter 3's question. Section 3.1 searched for a derivation of $P \to Q \vdash Q \to P$ and stalled. The row $p =$ F, $q =$ T makes $p \to q$ true (false hypothesis) and $q \to p$ false (true hypothesis, false conclusion). By Corollary 4.7.3,

$$p \to q \nvdash q \to p.$$

The letters of Chapter 3 stood for arbitrary statements, and the search asked for a derivation

of the **pattern**: one derivation, written with P and Q, that works whatever they are. Such a derivation, with P and Q read as the atoms p and q, would derive $p \to q \vdash q \to p$, which we have just ruled out. So the pattern has no derivation, and the search was right to fail. (Particular instances can be derivable: with P and Q the same statement, $P \to P \vdash P \to P$ takes one line. A refuting row for atoms speaks about the pattern, as the warning after Lemma 4.5.3 said.)

What soundness certifies. Soundness turns every break this book has made into a proof that no derivation exists. Section 1.1 called a form bad when some choice of statements makes its counterexample set true (Definition 1.1.3). For a propositional form, a row making the counterexample set true is such a choice, and by Corollary 4.7.3 no derivation of the form exists, with any rule of Chapters 1 and 2. Chapter 1 promised exactly this, in Section 1.3, for affirming the consequent: the row $p = \mathrm{F}$, $q = \mathrm{T}$ of Example (Affirming the consequent is not valid) makes $p \to q$ and q true and p false, so

$$p \to q,\ q \nvdash p.$$

So the repair of Argument B that Section 1.4 considered is not merely unjustified: it is impossible. The same holds for every break in the derive-or-break items from Section 1.5 on, for each claim of Section 3.2 that the search of Section 4.6 refuted, the four breaks of Exercise (B3: Derive or break, used later) among them, and for every refuting row of Sections 4.5 and 4.6. A break by statements about a number certifies the same: the truth values of those statements form a row, and since the connectives are truth-functional (Section 4.4), that row makes the counterexample set true.

Warning. A refuting row must make **every** premise true. For $p \to q,\ p \vdash q$, the row $p = \mathrm{F}$, $q = \mathrm{F}$ makes the conclusion false, but it also makes the premise p false, so it refutes nothing; and the sequent is derivable, by one line of $\to$E. A row in which some premise fails is simply not one of the rows Definition 4.5.3 looks at.

Warning. Soundness goes one way. It says that derivable sequents are valid; it does not say that a formula true in the row you happened to check is derivable. The atom p is true under the all-T valuation, but $\nvdash p$: the all-F valuation makes it false, so Corollary 4.7.3 (with Γ empty) applies. Whether **every** valid sequent is derivable is a separate question, answered in Section 4.8.

Quick check. Give a row showing $p \to r \nvdash (p \vee q) \to r$.

Solution. Take $p = \mathrm{F}$, $q = \mathrm{T}$, $r = \mathrm{F}$. The premise $p \to r$ is true, since p is false. The conclusion has hypothesis $p \vee q$, which is true, and conclusion r, which is false, so it is false. By Corollary 4.7.3, no derivation exists. (The rows with p true fail: they make the premise force r true, or make the premise false.)

Satisfiable lists are consistent. Section 4.5 kept two words apart, and soundness joins them in one direction. A satisfiable list Δ is consistent (Definition 3.2.1): if $\Delta \vdash \bot$, soundness would give $\Delta \models \bot$, but a row making Δ true makes $\bot$ false. With Proposition 3.2.2 (b) this is Corollary 4.7.3 once more: a row making the counterexample set $\Gamma, \neg\psi$ true shows it consistent, and so $\Gamma \nvdash \psi$, as Exercise B4 of Section 3.2 anticipated. Section 4.8 proves the converse.

4.7.7 Test the subgoal

Section 3.1's search works on subgoals: a goal, and the lines available for it. Soundness gives it a test. Before chasing a subgoal, look for a row making the available lines true and the subgoal false. If there is one, no line standing at this point can hold the subgoal: its context, the premises and the open assumptions, is among the available lines, so the row makes that context true and the subgoal false, while the proof of Theorem 4.7.2 showed that every line k satisfies $\Gamma_k \models \varphi_k$. This is the move **test the subgoal**. Section 3.1 met it without the name ("Choosing a side of $\vee$ can leave a goal that does not follow"): from $p \vee q$, the goal $q \vee p$ tempts $\vee$I from q, and the row $p = \mathrm{T}$, $q = \mathrm{F}$ makes $p \vee q$ true and q false, so the search must split $p \vee q$ first.

Chapter 2 met a harder case. Section 2.5 did not aim at $\vee$I for the De Morgan direction $\neg(P \wedge Q) \vdash \neg P \vee \neg Q$, because the step from the premise to either disjunct can be broken. Now such a break is a proof: the row $p = \mathrm{T}$, $q = \mathrm{F}$ makes $\neg(p \wedge q)$ true and $\neg p$ false, so $\neg(p \wedge q) \nvdash \neg p$, and the row $p = \mathrm{F}$, $q = \mathrm{T}$ gives $\neg(p \wedge q) \nvdash \neg q$. Both subgoals fail the test, so a derivation of Proposition 2.5.5 (d) cannot end with $\vee$I, which is why Section 2.5 reached for RAA. The test only refutes. A subgoal that passes it may still be hard to reach; whether it can be reached at all is Section 4.8's question.

Quick check. A search for $p \to q,\ \neg q \vdash \neg p$ opens a box p for $\neg$I, with the goal $\bot$. Does the subgoal q pass the test?

Solution. Yes, trivially. The available lines $p \to q$, $\neg q$, p are made true by no row at all (p and $p \to q$ force q), so no row refutes any subgoal there; and indeed q follows by $\to$E, then $\bot$ by $\neg$E. An unsatisfiable context lets every subgoal pass.

Remark (Looking ahead). Chapter 8 says what reductio asserts that the other rules do not, and who objected to it.

4.7.8 Exercises

4.7.8.1 A. Check your understanding

Exercise. A1

These parts are about the proof of Theorem 4.7.2. Part (c) uses the derivation of Proposition 2.1.2, written with atoms:

1	$\neg p$	
2	$\quad p$	asm, for $\to$I
3	$\quad \bot$	$\neg$E 2, 1
4	$\quad q$	$\bot$E 3
5	$p \to q$	$\to$I 2–4

(a) Why does the proof use strong induction on k, rather than ordinary induction from line k to line $k+1$?

(b) In the $\vee$E case, the induction hypothesis is applied to the last lines b and d of the two boxes. What does it say about each, and which part of Lemma 4.7.1 identifies the contexts?

(c) List the context of each of the five lines of the derivation above, and check $\Gamma_k \models \varphi_k$ for each.

(d) Suppose a rule "from $\varphi \to \psi$ and ψ, infer φ" were added to the primitive rules. Which part of the proof would fail? Give a valuation showing it.

(e) Where does Step 4 use that the last line of a derivation lies in no box?

Solution. *(a)* A rule may cite any available line, however far back, and the last line of a box that closed long ago. The hypothesis must therefore be available for every earlier line, not just for line $k-1$. For example, line 7 of the derivation of Proposition 2.3.2 uses the claim for line 6, and line 6 uses it for line 5; but line 5 uses it for line 2, three lines back.

(b) It says $\Gamma_b \models \chi$ and $\Gamma_d \models \chi$. By Lemma 4.7.1 (b), $\Gamma_b = \Gamma_k, \varphi$ and $\Gamma_d = \Gamma_k, \psi$: every row making Γ_k and φ true makes χ true, and every row making Γ_k and ψ true makes χ true. It is **not** $\Gamma_k \models \chi$, which is what we are trying to prove.

(c) The box is 2–4.

Line	Context	Check
1	$\neg p$	premise, in its context
2	$\neg p,\ p$	assumption, in its context
3	$\neg p,\ p$	no row makes $\neg p$ and p true, so it entails $\bot$
4	$\neg p,\ p$	the same unsatisfiable context entails q
5	$\neg p$	a row making $\neg p$ true makes p false, so $p \to q$ true

Line 3 is the $\neg$E case, line 4 the $\bot$E case, and line 5 the $\to$I case, with $\Gamma_4 = \Gamma_5, p$ by Lemma 4.7.1 (b).

(d) Claim 2, the one-step rules: it would need a new case, and the case is false. From $v(\varphi \to \psi) = \mathrm{T}$ and $v(\psi) = \mathrm{T}$ nothing follows about $v(\varphi)$. The valuation with $v(p) = \mathrm{F}$, $v(q) = \mathrm{T}$ makes $p \to q$ and q true and p false; with the new rule, the three-line derivation of p from $p \to q$ and q would have a last line whose context does not entail it.

(e) It concludes that the context of the last line is just the premises. Without that, the last line could depend on an undischarged assumption, and $\Gamma_n \models \psi$ would say nothing about the rows making the premises true.

Exercise. A2: Drill

The parts run from reading off a refuting row, through contexts, to testing a subgoal and

deciding a claim.

(a) Give a refuting row for $p \vee q \vdash p$.
(b) Give one for $p \to q,\ \neg p \vdash \neg q$.
(c) Does $p = \mathrm{T}$, $q = \mathrm{F}$ refute $p \wedge q \vdash q$?
(d) Is there a refuting row for $p,\ p \to q \vdash q$?
(e) Refute $\vdash p \vee q$.
(f) Refute $p \to (q \vee r) \vdash p \to q$.
(g) Line 5 of a derivation lies in the boxes opened at lines 2 and 4, and line 1 is the only premise. What is Γ_5?
(h) Line k is justified by $\neg$I citing the box 3–6. What is Γ_6, in terms of Γ_k?
(i) A search for $p \vee q \vdash q \vee p$ considers ending with $\vee$I from q. Test the subgoal q.
(j) The list $p \to q,\ q \to \neg p$ is satisfiable: give a row, and say what follows about $\vdash$.
(k) Which clause of Definition 4.4.2 does the $\vee$E case of the proof of Theorem 4.7.2 use?
(l) Derive or refute: $(p \wedge q) \to r \vdash (p \to r) \wedge (q \to r)$.

Solution.

(a) $p = \mathrm{F}$, $q = \mathrm{T}$.
(b) $p = \mathrm{F}$, $q = \mathrm{T}$.
(c) No: it makes the premise false.
(d) No. The only row making both premises true has q true. (The sequent is derivable, by one line of $\to$E.)
(e) $p = q = \mathrm{F}$. With no premises, every row qualifies.
(f) $p = \mathrm{T}$, $q = \mathrm{F}$, $r = \mathrm{T}$.
(g) $\varphi_1, \varphi_2, \varphi_4$.
(h) Γ_k, φ_3, by Lemma 4.7.1 (b).
(i) The row $p = \mathrm{T}$, $q = \mathrm{F}$ makes $p \vee q$ true and q false: $p \vee q \nvdash q$ (Corollary 4.7.3), so the subgoal is dead.
(j) $p = \mathrm{F}$, $q = \mathrm{T}$ (or both F). So the list is consistent: $p \to q,\ q \to \neg p \nvdash \bot$. (It is the list of Exercise (B4: Derive $\bot$, or make every member true, used later) (a) without its third member p; with p, Section 3.2 derived $\bot$.)
(k) (V4).
(l) Refute: $p = \mathrm{T}$, $q = \mathrm{F}$, $r = \mathrm{F}$ makes the premise true (false hypothesis) and $p \to r$ false, so the conclusion false; by Corollary 4.7.3 no derivation exists.

4.7.8.2 B. Practice

Exercise. B1: Derivable or not

Determine which of the following are derivable. Justify your answer.

(a) $(p \wedge q) \to r \vdash p \to r$.
(b) $\neg(p \wedge q) \vdash \neg p$.
(c) $p \leftrightarrow q,\ q \vee r \vdash p$.

(d) $p \to (q \vee r) \vdash (p \to q) \vee (p \to r)$.

Solution. *(a)* Not derivable. The row $p = \mathrm{T}$, $q = \mathrm{F}$, $r = \mathrm{F}$ makes the premise true (its hypothesis $p \wedge q$ is false) and the conclusion $p \to r$ false. By Corollary 4.7.3 there is no derivation.

(b) Not derivable. The row $p = \mathrm{T}$, $q = \mathrm{F}$ makes $p \wedge q$ false, so $\neg(p \wedge q)$ true, and $\neg p$ false. By Corollary 4.7.3 there is no derivation.

(c) Not derivable. The row $p = \mathrm{F}$, $q = \mathrm{F}$, $r = \mathrm{T}$ makes $p \leftrightarrow q$ true (both sides false) and $q \vee r$ true, and p false.

(d) Derivable: Example (Choosing a side too early) in Section 3.1 derives it, with LEM on q. So no row can refute it: by Theorem 4.7.2 it is valid. A search for a refuting row confirms this. The conclusion is false only if both $p \to q$ and $p \to r$ are false, that is, $p = \mathrm{T}$ and $q = r = \mathrm{F}$; but then $q \vee r$ is false and the premise is false.

The trap is (d). Its conclusion **looks** stronger than the premise, since it seems to decide in advance which disjunct follows from p; but a derivation exists, and trying to refute it by a row is bound to fail.

Exercise. B2: Ordering the $\vee$E case

Here are nine lines. Seven of them, in the right order, form the $\vee$E case of the proof of Theorem 4.7.2. Line k is χ, by $\vee$E citing line i (holding $\varphi \vee \psi$), the box from line a to line b (opening with φ, ending with χ), and the box from line c to line d (opening with ψ, ending with χ). Give an admissible order, and say why each of the other two lines cannot be used.

(i) By the induction hypothesis for lines b and d, which come before line k, $\Gamma_b \models \chi$ and $\Gamma_d \models \chi$.

(ii) Hence $v(\chi) = \mathrm{T}$ in either case; since v was arbitrary, $\Gamma_k \models \chi$.

(iii) Let v be a valuation making every formula of Γ_k true.

(iv) If $v(\psi) = \mathrm{T}$, then v makes every formula of Γ_d true, so $v(\chi) = \mathrm{T}$.

(v) By Lemma 4.7.1 (a), $\Gamma_i \subseteq \Gamma_k$, so v makes every formula of Γ_i true; by the induction hypothesis for line i, $v(\varphi \vee \psi) = \mathrm{T}$, so $v(\varphi) = \mathrm{T}$ or $v(\psi) = \mathrm{T}$.

(vi) By the induction hypothesis for line k, $\Gamma_k \models \chi$.

(vii) By Lemma 4.7.1 (b), Γ_b is Γ_k with φ added, and Γ_d is Γ_k with ψ added.

(viii) If $v(\varphi) = \mathrm{T}$, then v makes every formula of Γ_b true, so $v(\chi) = \mathrm{T}$.

(ix) Line b is inside the box, and the box was cited at line k, so the context of line b is Γ_k; the induction hypothesis for line b then gives $\Gamma_k \models \chi$ directly.

Solution. An admissible order: (iii), (v), (vii), (i), (viii), (iv), (ii).

What each line needs: (v) and (viii), (iv) use the valuation fixed in (iii); (viii) and (iv) use (vii) and (i); (ii) uses the case split of (v) and both of (viii) and (iv). So any order is admissible in which (iii) comes before (v), (viii), and (iv); (vii) and (i) come before (viii) and (iv); (v) comes before (ii); and (viii) and (iv) come before (ii). For instance (vii), (i), (iii), (v), (iv), (viii), (ii) is also admissible: (vii) and (i) do not mention v, and the two cases may

be taken in either order.

Line (vi) applies the induction hypothesis to line k itself. The hypothesis of Theorem 4.2.2 covers only the lines $j < k$; (vi) assumes what is to be proved.

Line (ix) uses the context of the citing line for the box's last line. It is false: line b lies inside the box, whose assumption φ is part of its context, so $\Gamma_b = \Gamma_k, \varphi$, not Γ_k. In the derivation of Example (Running the claim on a case split), line 3 has context $p \vee q,\ p$, while line 6 has context $p \vee q$, and $p \vee q \models q \vee p$ could not be read off from line 3 directly.

Exercise. B3: Derive or refute

For each claim, give a derivation, or a refuting row and the result that makes it a proof of non-derivability.

(a) $p \leftrightarrow q \vdash \neg p \leftrightarrow \neg q$.
(b) $p \leftrightarrow \neg q \vdash \neg(p \wedge q)$.
(c) $p \to q,\ q \to r \vdash r \to p$.
(d) $\neg p \vee q \vdash p \to q$.
(e) $p \leftrightarrow q,\ \neg q \vdash p$.
(f) $\neg(p \to q) \vdash q \to p$.

Solution.

(a) **Derive.**

1	$p \leftrightarrow q$	
2	$\quad \neg p$	asm, for $\leftrightarrow$I
3	$\quad\quad q$	asm, for $\neg$I
4	$\quad\quad p$	$\leftrightarrow$E 1, 3
5	$\quad\quad \bot$	$\neg$E 4, 2
6	$\quad \neg q$	$\neg$I 3–5
7	$\quad \neg q$	asm, for $\leftrightarrow$I
8	$\quad\quad p$	asm, for $\neg$I
9	$\quad\quad q$	$\leftrightarrow$E 1, 8
10	$\quad\quad \bot$	$\neg$E 9, 7
11	$\quad \neg p$	$\neg$I 8–10
12	$\neg p \leftrightarrow \neg q$	$\leftrightarrow$I 2–6, 7–11

(b) **Derive.**

1	$p \leftrightarrow \neg q$	
2	$\quad p \wedge q$	asm, for $\neg$I
3	$\quad p$	$\wedge$E 2
4	$\quad q$	$\wedge$E 2
5	$\quad \neg q$	$\leftrightarrow$E 1, 3
6	$\quad \bot$	$\neg$E 4, 5
7	$\neg(p \wedge q)$	$\neg$I 2–6

(c) **Refute.** $p = \mathrm{F}$, $q = \mathrm{F}$, $r = \mathrm{T}$: both premises have false hypotheses, and $r \to p$ is $\mathrm{T} \to \mathrm{F}$. By Corollary 4.7.3, no derivation exists.

(d) **Derive.**

1	$\neg p \vee q$	
2	$\quad p$	asm, for $\to$I
3	$\quad\quad \neg p$	asm, for $\vee$E
4	$\quad\quad \bot$	$\neg$E 2, 3
5	$\quad\quad q$	$\bot$E 4
6	$\quad\quad q$	asm, for $\vee$E
7	$\quad q$	$\vee$E 1, 3–5, 6–6
8	$p \to q$	$\to$I 2–7

(e) **Refute.** $p = q = \mathrm{F}$: $p \leftrightarrow q$ is T, $\neg q$ is T, and p is F. By Corollary 4.7.3, no derivation exists.

(f) **Derive**, without RAA.

1	$\neg(p \to q)$	
2	$\quad q$	asm, for $\to$I
3	$\quad\quad p$	asm, for $\to$I
4	$\quad\quad q$	R 2
5	$\quad p \to q$	$\to$I 3–4
6	$\quad \bot$	$\neg$E 5, 1
7	$\quad p$	$\bot$E 6
8	$q \to p$	$\to$I 2–7

(c) is the converse of the chain, a way of affirming a conclusion; (e) is a look-alike of $p \leftrightarrow q$, $\neg q \vdash \neg p$, which (a) gives by $\leftrightarrow$E; (f) looks paradoxical and is not: its premise is false in every row but one, and there $q \to p$ is true.

Exercise. B4: Test the subgoal

A search for $p \to (q \vee r),\ \neg q \vdash p \to r$ has reached this stage:

1	$p \to (q \vee r)$	
2	$\neg q$	
3	$\quad p$	asm, for $\to$I
	$\quad \vdots$	
9	$\quad r$	
10	$p \to r$	$\to$I 3–9

A student plans to reach r by deriving q first.

(a) Test that subgoal.

(b) *Name the move that works, then complete it.*

Solution.

(a) The available lines are $p \to (q \vee r)$, $\neg q$, p, and the subgoal is q. The row $p = \mathrm{T}$, $q = \mathrm{F}$, $r = \mathrm{T}$ makes the three lines true and q false. So no line of the box itself, whose context is $p \to (q \vee r)$, $\neg q$, p, can hold q: every line satisfies $\Gamma_k \models \varphi_k$, by the proof of Theorem 4.7.2. The plan is dead. (q may still stand as the assumption of a case box inside it, whose lines have a larger context; that is how (b) uses it.)

(b) Use the premise by its main connective: $\to$E gives $q \vee r$; split it, closing the case q by the clash with $\neg q$.

1	$p \to (q \vee r)$	
2	$\neg q$	
3	$\quad p$	asm, for $\to$I
4	$\quad q \vee r$	$\to$E 1, 3
5	$\quad\quad q$	asm, for $\vee$E
6	$\quad\quad \bot$	$\neg$E 5, 2
7	$\quad\quad r$	$\bot$E 6
8	$\quad\quad r$	asm, for $\vee$E
9	$\quad r$	$\vee$E 4, 5–7, 8–8
10	$p \to r$	$\to$I 3–9

(Proposition 2.3.4 gives line 9 from lines 4 and 2 in one step, as a derived rule; either is correct.)

4.7.8.3 C. Going deeper

Exercise. C1: A safe rule and an unsafe one

(a) Add to the primitive rules a one-step rule X: from a line $\varphi \wedge \psi$, infer $\varphi \vee \psi$. Show that Theorem 4.7.2 still holds for derivations that may use X.

(b) Add instead a one-step rule Y: from a line $\varphi \vee \psi$, infer $\varphi \wedge \psi$. Show that soundness fails: name the case of the proof that breaks, give a derivation using Y of a sequent with a refuting row, and give the row.

Solution (Hint 1). The proof of Theorem 4.7.2 treats one rule per case. Adding a rule adds one case to Claim 2 and changes nothing else.

Solution (Hint 2). For (b), start a derivation from the premise p, and look for a way to write $p \vee q$ first.

Solution. (a) The proof of Theorem 4.7.2 goes through with one more case in Claim 2. Every other part is unchanged: Step 1 and Step 4 do not look at the rules, Lemma 4.7.1 (a) holds for any cited line, and (b) concerns only box rules, and X is not one. *The case* X. Line k is $\varphi \vee \psi$, citing a line i holding $\varphi \wedge \psi$. By $(*)$, $v(\varphi \wedge \psi) = \mathrm{T}$, so $v(\varphi) = \mathrm{T}$ by (V3), and $v(\varphi \vee \psi) = \mathrm{T}$ by (V4). So every line of a derivation with X still satisfies $\Gamma_k \models \varphi_k$, and soundness holds.

(b) The new case of Claim 2 fails: from $v(\varphi \vee \psi) = \mathrm{T}$, (V4) gives only that one disjunct is true, and (V3) needs both. With Y, this is a derivation of $p \vdash q$:

1	p	
2	$p \vee q$	$\vee$I 1
3	$p \wedge q$	Y 2
4	q	$\wedge$E 3

The row $p = \mathrm{T}$, $q = \mathrm{F}$ makes the premise true and the conclusion false, so $p \not\models q$, although $p \vdash q$ with Y. The failure is at line 3: its context is p, and the row makes p and line 2 true but line 3 false.

Marking scheme. (10 points)

- 2: (a) says that only one case is added, and why the rest of the proof (Steps 1 and 4, Lemma 4.7.1) is unaffected.
- 3: (a) the case for X, using the hypothesis for the cited line through Lemma 4.7.1 (a) and the clauses (V3), (V4).
- 2: (b) names the failing case, with the reason (one true disjunct does not give a true conjunction).
- 3: (b) a correct derivation with Y and a refuting row for its sequent.

Fatal errors (0 for the step): a "refuting row" that makes a premise false; in (a), checking X only on atoms without the context argument.

Exercise. C2: Exit check: soundness

Close the section first, and write from memory.

Write out, in full, the →I, ¬I and ∨E cases of the proof of Theorem 4.7.2: state the claim being proved for every line, set up the induction, and prove the three cases, naming the reason for each step.

Solution. The claim. Fix a derivation in primitive rules with lines $1, \dots, n$, and write φ_k for the formula on line k and Γ_k for its context (the premises and the assumptions of the boxes containing line k). We prove $\Gamma_k \models \varphi_k$ for every k with $1 \le k \le n$, by Theorem 4.2.2 applied to "if $1 \le k \le n$, then $\Gamma_k \models \varphi_k$" (for $k = 0$ it holds vacuously): fix k with $1 \le k \le n$, assume $\Gamma_j \models \varphi_j$ for every line $j < k$, and let v be a valuation making every formula of Γ_k true. Every line and box that line k cites lies before line k, so the hypothesis applies to the cited lines and to the last lines of the cited boxes. We show $v(\varphi_k) = \mathrm{T}$.

→I. Line k is $\varphi \to \psi$, citing the box from line a (the assumption φ) to line b (the formula ψ). By Lemma 4.7.1 (b), $\Gamma_b = \Gamma_k, \varphi$. Since $b < k$, the hypothesis gives $\Gamma_k, \varphi \models \psi$. If $v(\varphi) = \mathrm{F}$, then $v(\varphi \to \psi) = \mathrm{T}$ by (V5). If $v(\varphi) = \mathrm{T}$, then v makes every formula of Γ_k, φ true, so $v(\psi) = \mathrm{T}$, and $v(\varphi \to \psi) = \mathrm{T}$ by (V5).

¬I. Line k is $\neg\varphi$, citing the box from line a (the assumption φ) to line b (the formula $\bot$). By Lemma 4.7.1 (b) and the hypothesis for line b, $\Gamma_k, \varphi \models \bot$. Since $v(\bot) = \mathrm{F}$ under every valuation (V1), no valuation makes all of Γ_k, φ true. As v makes Γ_k true, $v(\varphi) = \mathrm{F}$, so $v(\neg\varphi) = \mathrm{T}$ by (V2).

∨E. Line k is χ, citing line i (holding $\varphi \vee \psi$) and the boxes from a to b (φ to χ) and from c to d (ψ to χ). Line i is available at line k, so $\Gamma_i \subseteq \Gamma_k$ by Lemma 4.7.1 (a); v makes Γ_i true, and the hypothesis for line i gives $v(\varphi \vee \psi) = \mathrm{T}$, so $v(\varphi) = \mathrm{T}$ or $v(\psi) = \mathrm{T}$ by (V4). By Lemma 4.7.1 (b), $\Gamma_b = \Gamma_k, \varphi$ and $\Gamma_d = \Gamma_k, \psi$, and the hypothesis for lines b and d gives $\Gamma_k, \varphi \models \chi$ and $\Gamma_k, \psi \models \chi$. If $v(\varphi) = \mathrm{T}$, then v makes Γ_k, φ true, so $v(\chi) = \mathrm{T}$; if $v(\psi) = \mathrm{T}$, likewise by the second box.

Conclusion. With the other cases, the induction gives $\Gamma_k \models \varphi_k$ for every line. The last line lies in no box, so its context is the premises, which all belong to Γ; a valuation making Γ true makes the premises true, and so makes the last line ψ true: $\Gamma \models \psi$.

Marking scheme. (10 points)

- 2: states the claim $\Gamma_k \models \varphi_k$ for **every** line, with contexts.
- 2: uses the course-of-values hypothesis on earlier lines only, and says why the cited lines and boxes are earlier.
- 2: identifies the context of each box's last line as Γ_k plus the box's assumption, citing Lemma 4.7.1 (b) (and (a) for the cited line in ∨E).
- 1 + 1 + 1: the semantic step in each of the three cases (the rows where the assumption is false for →I; no row satisfies the box's context for ¬I; the split on which disjunct is true for ∨E).
- 1: the conclusion for the last line.

A proof that inducts on the length of the derivation instead, and handles open boxes correctly, earns the same points step for step. Fatal errors (0 for the step): applying the hypothesis to line k itself; using Γ_k as the context of a box's last line.

Exercise. C3: A changed table

Replace (V4) by: $v(\varphi \vee \psi) = \mathrm{T}$ exactly when **exactly one** of $v(\varphi)$, $v(\psi)$ is T. Keep every other clause and every rule.

(a) Which cases of the proof of Theorem 4.7.2 still go through?

(b) For each rule whose case fails, give a derivation and a row that is a counterexample for the new table.

(c) Is $p \vdash p \vee q$ still derivable? Is it valid for the new table? What does this show?

Solution.

(a) Every case except $\vee$I. The rules that do not mention $\vee$ use other clauses, which are unchanged. The $\vee$E case needs only that a true disjunction has a true side, and that still holds: exactly one side is true.

(b) $\vee$I: the derivation 1 p; 2 $p \vee q$, $\vee$I 1. The row $p = q = \mathrm{T}$ makes p true and, for the new table, $p \vee q$ false.

(c) Derivable: the rules did not change. Not valid for the new table: the row of (b). So soundness is a fact about the rules **and** the tables together; change a table, and the same rules can lead from truth to falsehood.

Marking scheme. (10 points)

- 4: (a), with the argument for the $\vee$E case.
- 4: (b), the derivation and the row, checked.
- 2: (c), both answers and the conclusion.

Fatal errors (0 for the step): in (a), declaring the $\vee$E case broken without a row that breaks it.

Exercise. C4: Challenge: $\bot$ read as a letter

Read $\bot$ as an ordinary letter. A **letter valuation** w gives T or F to every atom **and to** $\bot$, computes $\wedge$, $\vee$, $\to$, $\leftrightarrow$ by (V3)–(V6), and sets $w(\neg\varphi) = \mathrm{T}$ exactly when $w(\varphi) = \mathrm{F}$ or $w(\bot) = \mathrm{T}$. Write $\Gamma \models_\ell \psi$ when every letter valuation making every formula of Γ true makes ψ true.

(a) Show that every line of a derivation that uses no $\bot$E and no RAA, and cites no derived rule whose own derivation uses them, satisfies $\Gamma_k \models_\ell \varphi_k$. Give the cases $\neg$E and $\neg$I, and say why the others are those of Theorem 4.7.2.

(b) Deduce that $\bot \vdash p$ has no derivation without $\bot$E and RAA (and without derived rules derived with them).

(c) Derive $\bot \vdash p$ with RAA and without $\bot$E, and say why (b) must exclude RAA as well.

Solution.

(a) Step 1 of the proof expands each line citing a derived rule into a pasted derivation of that rule, which by hypothesis uses neither $\bot$E nor RAA; so the expanded derivation is in the remaining primitive rules (and a derived rule derived from derived rules of

this kind expands the same way, one layer at a time). Steps 1, 2 and 4 do not look at the clauses, and the cases of premises, assumptions, R and the rules for $\wedge$, $\vee$, $\to$, $\leftrightarrow$ use only (V3)–(V6), which are unchanged; they go through word for word with w for v.
$\neg$E: w makes Γ_k true, so by $(*)$ it makes φ and $\neg\varphi$ true. Now $w(\neg\varphi) = \mathrm{T}$ with $w(\varphi) = \mathrm{T}$ forces $w(\bot) = \mathrm{T}$, and line k is $\bot$.
$\neg$I: by $(**)$, every letter valuation making Γ_k and φ true makes $\bot$ true. If $w(\varphi) = \mathrm{F}$, then $w(\neg\varphi) = \mathrm{T}$. If $w(\varphi) = \mathrm{T}$, then w makes Γ_k and φ true, so $w(\bot) = \mathrm{T}$, and again $w(\neg\varphi) = \mathrm{T}$.

(b) The letter valuation with $w(\bot) = \mathrm{T}$ and $w(p) = \mathrm{F}$ makes the premise true and the conclusion false. So $\bot \models_\ell p$ fails, and by (a) no derivation of $\bot \vdash p$ without $\bot$E and RAA, and without derived rules derived with them, exists. (Many derived rules of Chapter 2 use one of the two, Proposition 2.5.1 (b) (with RAA) and Proposition 2.1.1 (with $\bot$E) among them; a derivation that avoids the two only on its face is not covered.)

(c)

1	$\bot$	
2	$\quad\neg p$	asm, for RAA
3	$\quad\bot$	R 1
4	p	RAA 2–3

So with RAA the rule $\bot$E is derivable, and (a) must fail for RAA. It does: with $w(\bot) = \mathrm{T}$, every negation is true, so the box of lines 2–3 satisfies $(**)$, and yet $w(p) = \mathrm{F}$.

Marking scheme. (10 points)

- 2: the carried-over cases named, with the reason (Steps 1, 2, 4 and the clauses (V3)–(V6) unchanged), and Step 1's restriction to derived rules derived without $\bot$E and RAA.
- 2: the $\neg$E case.
- 3: the $\neg$I case, both values of $w(\varphi)$.
- 1: (b), the letter valuation and the conclusion.
- 2: (c), the derivation and the failing case.

Fatal errors (0 for the step): in (b), using the ordinary valuations, under which $\bot \models p$ holds.

Exercise. C5: Challenge: two three-valued tables

A **Kleene valuation** u_{K} assigns one of the values 0, $\frac{1}{2}$, 1 to each atom and is extended by $u_{\mathrm{K}}(\bot) = 0$, $u_{\mathrm{K}}(\neg\varphi) = 1 - u_{\mathrm{K}}(\varphi)$, $u_{\mathrm{K}}(\varphi \wedge \psi) = \min\{u_{\mathrm{K}}(\varphi), u_{\mathrm{K}}(\psi)\}$, $u_{\mathrm{K}}(\varphi \vee \psi) = \max\{u_{\mathrm{K}}(\varphi), u_{\mathrm{K}}(\psi)\}$, $u_{\mathrm{K}}(\varphi \to \psi) = \max\{1 - u_{\mathrm{K}}(\varphi), u_{\mathrm{K}}(\psi)\}$. (This extension differs from the three-valued valuation of the optional excursion below, whose $\neg$ gives $\neg\varphi$ the value 0 at $\frac{1}{2}$.) For a set D of **designated** values, say that Γ D-entails ψ if every Kleene valuation u_{K} giving every formula of Γ a value in D gives ψ a value in D.

(a) For $D = \{1\}$, show that the case $\to$I fails: give a box that is correct for D-entailment and a u_{K} at which the line that $\to$I writes is undesignated.

(b) For $D = \{\frac{1}{2}, 1\}$, show that the case $\to$E fails.

(c) Is $p \vee \neg p$ designated under every u_{K} in (a)? In (b)?

Solution.

(a) The one-line box p: p D-entails p. $\to$I writes $p \to p$, and at $u_{\mathrm{K}}(p) = \frac{1}{2}$, $u_{\mathrm{K}}(p \to p) = \max\{\frac{1}{2}, \frac{1}{2}\} = \frac{1}{2} \notin D$.

(b) The premises $p \to q$, p and the conclusion q, at $u_{\mathrm{K}}(p) = \frac{1}{2}$, $u_{\mathrm{K}}(q) = 0$: $u_{\mathrm{K}}(p \to q) = \max\{\frac{1}{2}, 0\} = \frac{1}{2} \in D$ and $u_{\mathrm{K}}(p) = \frac{1}{2} \in D$, but $u_{\mathrm{K}}(q) = 0 \notin D$.

(c) In (a), no: $u_{\mathrm{K}}(p \vee \neg p) = \max\{\frac{1}{2}, \frac{1}{2}\} = \frac{1}{2}$ at $u_{\mathrm{K}}(p) = \frac{1}{2}$. In (b), yes: $\max\{x, 1 - x\} \geq \frac{1}{2}$ for each of the three values x.

So the first choice of D loses excluded middle and the box rule for $\to$; the second keeps excluded middle and loses modus ponens.

Marking scheme. (10 points)

- 4: (a), the box, the valuation and the computation.
- 3: (b), the valuation and the three values.
- 3: (c), both parts, with the computation.

Fatal errors (0 for the step): a "failure" in (a) or (b) at a valuation that does not designate every premise.

4.7.9 Optional excursion: reductio says something new

Optional: nothing later depends on this excursion. It answers a question Chapter 2 raised.

A **three-valued valuation** u assigns to each atom one of the values 0, $\frac{1}{2}$, 1, ordered $0 < \frac{1}{2} < 1$. Think of 1 as "true", 0 as "false", and $\frac{1}{2}$ as "not yet settled".

$u(\varphi \to \psi)$	$u(\psi) = 0$	$u(\psi) = \frac{1}{2}$	$u(\psi) = 1$
$u(\varphi) = 0$	1	1	1
$u(\varphi) = \frac{1}{2}$	0	1	1
$u(\varphi) = 1$	0	$\frac{1}{2}$	1

$u(\varphi)$	$u(\neg\varphi)$
0	1
$\frac{1}{2}$	0
1	0

An implication gets 1 when its conclusion is at least as true as its hypothesis, and otherwise the value of its conclusion; a negation is 1 only on 0.

Chapter 2 introduced RAA in Section 2.5 because the rules before it stalled on $\neg\neg P \vdash P$, and said that Chapter 4 would prove that no derivation of P from $\neg\neg P$ avoids RAA. Truth tables cannot do it. The row method only proves non-derivability from **all** the rules, and with RAA the sequent is derivable, so no two-valued row refutes it. The trick is to run the proof of Theorem 4.7.2 with a different notion of truth, one with three values, under which every rule except RAA still passes.

A three-valued valuation is extended to all formulas by recursion (Theorem 4.3.4):

$$u(\bot) = 0, \qquad u(\neg\varphi) = u(\varphi \to \bot),$$
$$u(\varphi \wedge \psi) = \min\{u(\varphi), u(\psi)\}, \qquad u(\varphi \vee \psi) = \max\{u(\varphi), u(\psi)\},$$
$$u(\varphi \to \psi) = \begin{cases} 1 & \text{if } u(\varphi) \le u(\psi), \\ u(\psi) & \text{otherwise,} \end{cases}$$
$$u(\varphi \leftrightarrow \psi) = \min\{u(\varphi \to \psi), u(\psi \to \varphi)\}.$$

So $u(\neg\varphi) = 1$ if $u(\varphi) = 0$, and $u(\neg\varphi) = 0$ otherwise. On the values 0 and 1 alone these are exactly the tables of Definition 4.4.2, with 1 for T and 0 for F. For a list Γ, let $u(\Gamma)$ be the least of the values $u(\gamma)$ for γ in Γ, and $u(\Gamma) = 1$ if Γ is empty. Write $\Gamma \models_3 \psi$ if $u(\Gamma) \le u(\psi)$ for every three-valued valuation u: the conclusion is at least as true as the weakest premise.

Theorem 4.7.4. Reductio Is Not Derivable from the Other Rules

Optional: nothing later depends on this.

Neither $\neg\neg p \vdash p$ nor $\vdash p \vee \neg p$ has a derivation that uses only the primitive rules other than RAA (and derived rules derived from those).

Idea. Repeat the proof of Theorem 4.7.2 with $\models_3$ in place of $\models$: every line of a derivation without RAA satisfies $\Gamma_k \models_3 \varphi_k$. Then show that both sequents fail for $\models_3$, at $u(p) = \frac{1}{2}$. The one case that breaks is RAA, and that is the point.

Proof. As in Step 1 of the proof of Theorem 4.7.2, a derived rule derived without RAA abbreviates a pasted derivation in the primitive rules other than RAA, so it is enough to consider those. Fix a derivation that uses only the primitive rules other than RAA, with lines $1, \dots, n$, formulas φ_k and contexts Γ_k. We show $\Gamma_k \models_3 \varphi_k$ for every k with $1 \le k \le n$, by Theorem 4.2.2 exactly as in Step 2 of the proof of Theorem 4.7.2, the cases $k = 0$ and $k > n$ being vacuous. Fix k, assume the claim for all earlier lines, and fix a three-valued valuation u; write $m = u(\Gamma_k)$. We must show $m \le u(\varphi_k)$. Two facts do the bookkeeping. If line k cites line i, then $\Gamma_i \subseteq \Gamma_k$ (Lemma 4.7.1 (a)), and the least value over a longer list is at most the least over a shorter one, so $m \le u(\Gamma_i) \le u(\varphi_i)$. If line k cites a box from line a to line b by a box rule, then $\Gamma_b = \Gamma_k, \varphi_a$ (Lemma 4.7.1 (b)), so the hypothesis for line b says $\min\{m, u(\varphi_a)\} \le u(\varphi_b)$. Below, $x = u(\varphi)$, $y = u(\psi)$, $z = u(\chi)$.

Premises and assumptions: φ_k is in Γ_k, so $m \le u(\varphi_k)$. R*:* $m \le u(\varphi_i) = u(\varphi_k)$. $\wedge$I*:* $m \le x$ and $m \le y$, so $m \le \min\{x, y\}$. $\wedge$E*:* $m \le \min\{x, y\} \le x$, and likewise $\le y$. $\vee$I*:* $m \le x \le \max\{x, y\}$.

$\to$E*:* $m \le u(\varphi \to \psi)$ and $m \le x$. If $x \le y$, then $m \le x \le y$. Otherwise $u(\varphi \to \psi) = y$, so $m \le y$. $\leftrightarrow$E*:* $m \le u(\varphi \leftrightarrow \psi) \le u(\varphi \to \psi)$ and $m \le x$, so $m \le y$ as for $\to$E; the second form is symmetric.

$\neg$E*:* $m \le x$ and $m \le u(\neg\varphi)$. If $x = 0$, then $m \le 0$; if $x > 0$, then $u(\neg\varphi) = 0$ and $m \le 0$. So $m = 0 = u(\bot)$. $\bot$E*:* $m \le u(\bot) = 0$, so $m = 0 \le u(\varphi_k)$.

$\to$I*:* the hypothesis gives $\min\{m, x\} \le y$. If $x \le y$, then $u(\varphi \to \psi) = 1 \ge m$. Otherwise $y < x$ and $u(\varphi \to \psi) = y$; since $\min\{m, x\} \le y < x$, the minimum is not x, so it is m, and $m \le y$. $\leftrightarrow$I*:* the same argument for each box gives $m \le u(\varphi \to \psi)$ and $m \le u(\psi \to \varphi)$, so m

is at most their minimum.

$\neg$I*:* the hypothesis gives $\min\{m, x\} \leq u(\bot) = 0$. If $x = 0$, then $u(\neg\varphi) = 1 \geq m$. If $x > 0$, the minimum is 0 and is not x, so $m = 0 \leq u(\neg\varphi)$.

$\vee$E*:* $m \leq \max\{x, y\}$, and the boxes give $\min\{m, x\} \leq z$ and $\min\{m, y\} \leq z$. If $x \geq y$, then $m \leq x$, so $\min\{m, x\} = m$ and $m \leq z$. If $y > x$, then $m \leq y$ and $m \leq z$ from the second box.

This covers every rule used, so $\Gamma_k \models_3 \varphi_k$ for every line, and in particular for the last line, whose context is the premises. As in Step 4 of the proof of Theorem 4.7.2, a list containing all the premises has least value at most theirs, so every sequent $\Gamma \vdash \psi$ derived without RAA satisfies $\Gamma \models_3 \psi$.

Now let $u(p) = \frac{1}{2}$. Then $u(\neg p) = 0$, since $\frac{1}{2} > 0$, and $u(\neg\neg p) = 1$, since $u(\neg p) = 0$. So $u(\neg\neg p) = 1 > \frac{1}{2} = u(p)$, and $\neg\neg p \models_3 p$ fails. Also $u(p \vee \neg p) = \max\{\frac{1}{2}, 0\} = \frac{1}{2} < 1$, the value of the empty list, so $\models_3 p \vee \neg p$ fails. Hence neither sequent is derivable without RAA.

■

The case that was not written is RAA itself. Its box would give $\min\{m, u(\neg\varphi)\} \leq 0$, and we would need $m \leq x$. At $u(\varphi) = \frac{1}{2}$ the box gives nothing, since $u(\neg\varphi) = 0$ already, while m may be 1. This is the three-valued face of the question asked in the Idea of Theorem 4.7.2: the two-valued RAA case used that $v(\varphi) = \mathrm{F}$ forces $v(\neg\varphi) = \mathrm{T}$, and with a third value, φ and $\neg\varphi$ can both fail to be true.

Remark. The same values settle Peirce's law, $((p \to q) \to p) \to p$, which Section 2.5 derived with RAA (Proposition 2.5.4): at $u(p) = \frac{1}{2}$, $u(q) = 0$, we get $u(p \to q) = 0$, $u((p \to q) \to p) = 1$, and the whole formula has value $\frac{1}{2}$, so it too needs RAA. The values do not refute everything Chapter 2 derived with RAA: the De Morgan direction $\neg(p \wedge q) \vdash \neg p \vee \neg q$ passes this test, and so does $\vdash (p \to q) \vee (q \to p)$. Neither of these is derivable without RAA either, but showing that needs a different kind of model, and this book states it without proof, as Section 2.5's Looking-ahead box announced. RAA says something the other rules cannot.

A technique, named. The proof is an instance of a general technique: to show that a rule cannot be derived from the others, find a new notion of "true" under which the other rules are sound and that rule is not. Here the notion has three values; Exercise C4 uses two values with $\bot$ read as an ordinary letter, and Exercise C5 two other three-valued tables. Each notion is a different semantics. For the two-valued tables the rules are not only sound but complete, which is Section 4.8's theorem.

4.8 Completeness

Soundness (Theorem 4.7.2) says that the rules of Chapters 1 and 2 never lead from true premises to a false conclusion: whatever can be derived is true in every row where the premises are. That answers half of the chapter's question. The other half is the converse. Could the rules be too **weak**, so that some sequent is true in every row and yet no derivation of it exists? This section proves that they are not: every truth-table consequence has a derivation. The proof reads a derivation off the truth table, one row at a time, and then glues the rows together with excluded

middle.

4.8.1 The question soundness leaves open

Recall the two relations from Section 4.5. $\Gamma \vdash \psi$ says that **there is** a derivation of ψ from premises in Γ (Definition 1.2.3). $\Gamma \models \psi$ says that **every** valuation making all of Γ true makes ψ true (Definition 4.5.3). Soundness is the implication

$$\Gamma \vdash \psi \;\Rightarrow\; \Gamma \models \psi.$$

Its converse is not automatic. The rules are a list that people wrote down, and a list can leave something out. Suppose, as a thought experiment, that we struck RAA (Definition 2.5.1) from the list. Every derivation that remains is still a derivation, so soundness still holds. But the tautology $p \vee \neg p$ would then have no derivation (this is the optional Theorem 4.7.4 of Section 4.7). So a sound list of rules can be too short. The question is whether **our** list is long enough.

Here is a sequent to keep in mind. The formula $(p \to q) \vee (q \to p)$ is a tautology: if q is true then $p \to q$ is true, and if q is false then $q \to p$ is true. Chapter 3 found a derivation of it (Exercise (C2: A disjunction from nothing)), but only after reaching for LEM, and nothing in Chapter 3 proved that a derivation would be there to find. We want a theorem that promises it in advance, for every tautology and every valid sequent.

We are going to prove:

> If Γ is a list of formulas and $\Gamma \models \psi$, then $\Gamma \vdash \psi$.

The plan has two ideas, one per row and one for all the rows together.

1. **One row gives one derivation.** Fix a row of the truth table, that is, a valuation v. Write each atom as it stands in that row: p if v makes p true, $\neg p$ if v makes it false. From these signed atoms we will derive the formula ψ if ψ is true in the row, and $\neg\psi$ if it is false. The derivation climbs the parse tree of ψ, one connective at a time.
2. **The rows are glued by excluded middle.** If $\Gamma \models \psi$, the first idea gives ψ from Γ and the signed atoms of **each** row. Two rows that differ only at the atom p have derivations from "$\ldots, p$" and from "$\ldots, \neg p$". The rule $\vee$E applied to $p \vee \neg p$ merges them into one derivation that no longer mentions p. Removing the atoms one at a time leaves a derivation from Γ alone.

We keep the premises Γ throughout, and we isolate the two ingredients as lemmas so that each can be checked on its own. Before either, we need a little bookkeeping about derivations.

Second reading · *Where the route comes from*

The first idea is known as Kalmár's lemma, and gluing rows by excluded middle is the standard way to finish; the route is the classical one.

4.8.2 Bookkeeping: adding premises and chaining derivations

In Chapter 2 we learned to use a proved sequent as a single line (Definition 2.3.1), and Section 2.3 proved what that gives, Lemma 2.3.3: (a) extra premises never hurt; (b) a statement already

derived may be used as a premise; (c) a sequent already proved may be applied to statements already derived. The completeness proof uses (a) and (c) dozens of times. They also made Step 1 of the proof of Theorem 4.7.2 honest: for whole derivations, part (c) is the precise form of "abbreviates".

Quick check. Suppose $\Gamma \vdash \alpha$ and $\Gamma \vdash \neg\alpha$. Which part of Lemma 2.3.3, applied to which proved sequent, gives $\Gamma \vdash \chi$ for every formula χ?

Solution. Part (c) with $m = 2$, $\alpha_1 = \alpha$, $\alpha_2 = \neg\alpha$ and $\beta = \chi$, applied to Proposition 2.1.1, $P, \neg P \vdash Q$, with α for P and χ for Q.

4.8.3 The rules prove their own truth tables

Each line of the truth table of a connective says something like "with α true and β false, $\alpha \to \beta$ comes out false". Read with signs, that line is a sequent: from α and $\neg\beta$, conclude $\neg(\alpha \to \beta)$. If every such sequent is derivable, then the rules know the truth tables. This is the key to the first idea of the plan. Section 4.4 began by reading the tables off the rules; the lemma below makes that reading exact: every line of every table is a derivable sequent.

Lemma 4.8.1. The Rules Prove the Truth Tables

For all formulas α and β, the following sequents are derivable, by derivations that do not use RAA.

(i) $(\neg 1)$ $\alpha \vdash \neg\neg\alpha$, and $(\bot 1) \vdash \neg\bot$.

(ii) $(\wedge 1)$ $\alpha, \beta \vdash \alpha \wedge \beta$; $(\wedge 2)$ $\neg\alpha \vdash \neg(\alpha \wedge \beta)$; $(\wedge 3)$ $\neg\beta \vdash \neg(\alpha \wedge \beta)$.

(iii) $(\vee 1)$ $\alpha \vdash \alpha \vee \beta$; $(\vee 2)$ $\beta \vdash \alpha \vee \beta$; $(\vee 3)$ $\neg\alpha, \neg\beta \vdash \neg(\alpha \vee \beta)$.

(iv) $(\to 1)$ $\neg\alpha \vdash \alpha \to \beta$; $(\to 2)$ $\beta \vdash \alpha \to \beta$; $(\to 3)$ $\alpha, \neg\beta \vdash \neg(\alpha \to \beta)$.

(v) $(\leftrightarrow 1)$ $\alpha, \beta \vdash \alpha \leftrightarrow \beta$; $(\leftrightarrow 2)$ $\neg\alpha, \neg\beta \vdash \alpha \leftrightarrow \beta$; $(\leftrightarrow 3)$ $\alpha, \neg\beta \vdash \neg(\alpha \leftrightarrow \beta)$; $(\leftrightarrow 4)$ $\neg\alpha, \beta \vdash \neg(\alpha \leftrightarrow \beta)$.

Here are the tables of Definition 4.4.2 again, with each row tagged by the sequent of the lemma that derives it from the signed parts. Where two sequents fit a row, the tag is the one the proof of the next lemma uses.

α	β	$\alpha \wedge \beta$	$\alpha \vee \beta$	$\alpha \to \beta$	$\alpha \leftrightarrow \beta$
T	T	T $(\wedge 1)$	T $(\vee 1)$	T $(\to 2)$	T $(\leftrightarrow 1)$
T	F	F $(\wedge 3)$	T $(\vee 1)$	F $(\to 3)$	F $(\leftrightarrow 3)$
F	T	F $(\wedge 2)$	T $(\vee 2)$	T $(\to 1)$	F $(\leftrightarrow 4)$
F	F	F $(\wedge 2)$	F $(\vee 3)$	T $(\to 1)$	T $(\leftrightarrow 2)$

Beside it: α T gives $\neg\alpha$ F, tagged $(\neg 1)$; α F gives $\neg\alpha$ T, and no sequent is needed, since the signed part $\neg\alpha$ is already the signed whole; $\bot$ is always F, tagged $(\bot 1)$.

Each sequent has as premises **only the parts it needs**, each as itself or negated. For instance, $\alpha \wedge \beta$ is false as soon as α is false, whatever β is, and $(\wedge 2)$ has the single premise $\neg\alpha$. This

economy will matter: when the proof of the next lemma uses a row, it needs to know something only about the parts the row mentions.

Idea. Most of the sequents are already in the book: rules of Chapter 1, or results of Chapter 2 with α and β put in for the letters. The others are short derivations of the kinds Chapter 2 practiced: a box that repeats a premise, or a $\neg$I box that reaches a clash. The only thing to watch is that no step, and no cited result, uses RAA.

The derivations, with a check that none uses RAA, are in the fold below.

Second reading · *The fifteen derivations*

Proof. The table lists, for each sequent, the rows of the truth table it covers (by Definition 4.4.2) and where its derivation comes from. The letters of the Chapter 2 results are replaced by α and β, as Definition 2.3.1 allows.

Tag	Rows covered	Sequent	Derivation
($\neg$1)	α T, so $\neg\alpha$ F	$\alpha \vdash \neg\neg\alpha$	Example (Two negations from one premise) (a)
($\bot$1)	$\bot$ is always F	$\vdash \neg\bot$	printed below
($\wedge$1)	T, T give T	$\alpha, \beta \vdash \alpha \wedge \beta$	$\wedge$I (Definition 1.2.5)
($\wedge$2)	F, anything give F	$\neg\alpha \vdash \neg(\alpha \wedge \beta)$	$\vee$I gives $\neg\alpha \vee \neg\beta$, then Proposition 2.5.5 (c)
($\wedge$3)	anything, F give F	$\neg\beta \vdash \neg(\alpha \wedge \beta)$	$\vee$I gives $\neg\alpha \vee \neg\beta$, then Proposition 2.5.5 (c)
($\vee$1)	T, anything give T	$\alpha \vdash \alpha \vee \beta$	$\vee$I (Definition 1.5.1)
($\vee$2)	anything, T give T	$\beta \vdash \alpha \vee \beta$	$\vee$I
($\vee$3)	F, F give F	$\neg\alpha, \neg\beta \vdash \neg(\alpha \vee \beta)$	$\wedge$I gives $\neg\alpha \wedge \neg\beta$, then Proposition 2.5.5 (b)
($\to$1)	F, anything give T	$\neg\alpha \vdash \alpha \to \beta$	Proposition 2.1.2
($\to$2)	anything, T give T	$\beta \vdash \alpha \to \beta$	printed below
($\to$3)	T, F give F	$\alpha, \neg\beta \vdash \neg(\alpha \to \beta)$	printed below
($\leftrightarrow$1)	T, T give T	$\alpha, \beta \vdash \alpha \leftrightarrow \beta$	printed below
($\leftrightarrow$2)	F, F give T	$\neg\alpha, \neg\beta \vdash \alpha \leftrightarrow \beta$	printed below
($\leftrightarrow$3)	T, F give F	$\alpha, \neg\beta \vdash \neg(\alpha \leftrightarrow \beta)$	printed below
($\leftrightarrow$4)	F, T give F	$\neg\alpha, \beta \vdash \neg(\alpha \leftrightarrow \beta)$	printed below

In the second column, "F, anything give F" means: if α is F, then $\alpha \wedge \beta$ is F whatever the value of β.

The seven printed derivations follow. First ($\to$2) and ($\to$3) (the first is the derivation of Exercise (B2: An assumption that is not used) (a), printed again because a proof does not rest on an exercise): an implication whose conclusion we already have, and the refutation of an implication by a true hypothesis with a false conclusion.

1	β	
2	$\quad\alpha$	asm, for $\to$I
3	$\quad\beta$	R 1
4	$\alpha \to \beta$	$\to$I 2–3

1	α	
2	$\neg\beta$	
3	$\quad\alpha \to \beta$	asm, for $\neg$I
4	$\quad\beta$	$\to$E 3, 1
5	$\quad\bot$	$\neg$E 4, 2
6	$\neg(\alpha \to \beta)$	$\neg$I 3–5

Next ($\leftrightarrow$1) and ($\leftrightarrow$2). In the first, each box repeats the premise it needs; in the second, each box reaches a clash and closes by $\bot$E.

1	α	
2	β	
3	$\quad\alpha$	asm, for $\leftrightarrow$I
4	$\quad\beta$	R 2
5	$\quad\beta$	asm, for $\leftrightarrow$I
6	$\quad\alpha$	R 1
7	$\alpha \leftrightarrow \beta$	$\leftrightarrow$I 3–4, 5–6

1	$\neg\alpha$	
2	$\neg\beta$	
3	$\quad\alpha$	asm, for $\leftrightarrow$I
4	$\quad\bot$	$\neg$E 3, 1
5	$\quad\beta$	$\bot$E 4
6	$\quad\beta$	asm, for $\leftrightarrow$I
7	$\quad\bot$	$\neg$E 6, 2
8	$\quad\alpha$	$\bot$E 7
9	$\alpha \leftrightarrow \beta$	$\leftrightarrow$I 3–5, 6–8

Then ($\leftrightarrow$3), ($\leftrightarrow$4) and ($\bot$1). The first two assume the biconditional and use it in the direction that clashes with a premise; the last is a box whose assumption $\bot$ is already the $\bot$ that $\neg$I asks for, the degenerate box of Section 2.2.

1	α	
2	$\neg\beta$	
3	$\quad\alpha \leftrightarrow \beta$	asm, for $\neg$I
4	$\quad\beta$	$\leftrightarrow$E 3, 1
5	$\quad\bot$	$\neg$E 4, 2
6	$\neg(\alpha \leftrightarrow \beta)$	$\neg$I 3–5

1	$\neg\alpha$	
2	β	
3	$\quad\alpha \leftrightarrow \beta$	asm, for $\neg$I
4	$\quad\alpha$	$\leftrightarrow$E 3, 2
5	$\quad\bot$	$\neg$E 4, 1
6	$\neg(\alpha \leftrightarrow \beta)$	$\neg$I 3–5

1	$\quad\bot$	asm, for $\neg$I
2	$\neg\bot$	$\neg$I 1–1

None of these derivations uses RAA: their boxes close by →I, ¬I and ↔I. Nor do the cited results: Proposition 2.1.2 and Example (Two negations from one premise) come from Sections 2.1 and 2.2, before RAA was introduced, and Chapter 2 derived De Morgan's laws (b) and (c) without it (it says so right after the derivations of Proposition 2.5.5). This proves the lemma.

■

The ¬ connective needs only one sequent. When α is false, $\neg\alpha$ is true, and the "sequent" $\neg\alpha \vdash \neg\alpha$ is a one-line derivation; only the other row, (¬1), needs a proof.

Warning. The sequents of Lemma 4.8.1 go **from the parts to the whole**. Read backwards they are false. For instance $\alpha \vee \beta \vdash \alpha$ is not derivable in general: the row $p = \mathrm{F}$, $q = \mathrm{T}$ makes $p \vee q$ true and p false, so $p \vee q \nvdash p$ by Corollary 4.7.3. A row of a table tells us the value of the whole from the values of the parts, never the other way round.

Quick check. Which tagged sequent covers α F, β T for →? And for ↔?

Solution. (→1), $\neg\alpha \vdash \alpha \to \beta$. (The row also fits (→2).) For ↔: (↔4), $\neg\alpha, \beta \vdash \neg(\alpha \leftrightarrow \beta)$.

4.8.4 One row, one derivation

A row of a truth table can be written as a list of premises: each atom as it stands in the row, itself if true, negated if false.

p	q	r	the row, as premises
T	F	T	$p,\ \neg q,\ r$

We now name the formulas that "write down a row".

A signed formula is a formula as it stands in one row: itself if the row makes it true, its negation if the row makes it false.

For a valuation v and a formula φ, the **signed formula** φ^v is

$$\varphi^v := \begin{cases} \varphi & \text{if } v(\varphi) = \mathrm{T}, \\ \neg\varphi & \text{if } v(\varphi) = \mathrm{F}. \end{cases}$$

The superscript v is a label naming the row, not an exponent. In particular, for an atom p, the **signed atom** p^v is p or $\neg p$. For example, if $v(p) = \mathrm{T}$ and $v(q) = \mathrm{F}$, then $p^v = p$, $q^v = \neg q$, $(p \to q)^v = \neg(p \to q)$ and $(q \to p)^v = q \to p$. Whatever v is, $v(\varphi^v) = \mathrm{T}$: the signed formula is always the true one of φ and $\neg\varphi$, by the clause for ¬ in Definition 4.4.2.

The signed atoms $p_1^v, \ldots, p_n^v$ describe the row of v completely, as far as $p_1, \ldots, p_n$ go. The next

lemma says they are enough premises to derive the signed version of any formula built from these atoms.

Lemma 4.8.2. One Row of the Truth Table Is Derivable

Let φ be a formula, let $p_1, \ldots, p_n$ ($n \geq 0$) be atoms among which are **all** the atoms of φ, and let v be a valuation. Then

$$p_1^v, \ldots, p_n^v \vdash \varphi^v.$$

In words: in each row of the truth table of φ, the value of φ can be derived from the values of its atoms, where "value T" is written as the formula and "value F" as its negation.

Idea. Climb the parse tree. At an atom, φ^v is one of the premises. At a connective, the induction hypothesis gives the signed parts α^v, β^v, and Lemma 4.8.1 has a sequent from (some of) those signed parts to the signed whole. So structural induction (Theorem 4.3.1), with Lemma 2.3.3 (c) to chain, does the whole job. The only care is to pick, for each pair of values, the row whose premises are the signs we actually have.

Proof. Fix the atoms $p_1, \ldots, p_n$ and the valuation v, and write Σ for the list $p_1^v, \ldots, p_n^v$. We prove, by structural induction (Theorem 4.3.1), that every formula φ has the property

> if every atom of φ is among $p_1, \ldots, p_n$, then $\Sigma \vdash \varphi^v$.

Every atom of an immediate part α or β of a formula is an atom of the formula itself, so when φ satisfies the "if", so do its immediate parts, and the induction hypothesis applies to them.

Atoms. Let φ be an atom among $p_1, \ldots, p_n$, say $\varphi = p_i$. Then $\varphi^v = p_i^v$ is a member of Σ, and the one-line derivation whose only line is the premise p_i^v shows $\Sigma \vdash \varphi^v$.

The formula $\bot$. By Definition 4.4.2, $v(\bot) = \mathrm{F}$, so $\bot^v = \neg\bot$. By ($\bot$1) of Lemma 4.8.1, $\vdash \neg\bot$, and weakening (Lemma 2.3.3 (a)) gives $\Sigma \vdash \neg\bot$.

Negation. Let $\varphi = \neg\alpha$, and assume $\Sigma \vdash \alpha^v$.

- *Case* $v(\alpha) = \mathrm{T}$. Then $\alpha^v = \alpha$, and $v(\neg\alpha) = \mathrm{F}$, so $\varphi^v = \neg\neg\alpha$. From $\Sigma \vdash \alpha$ and ($\neg$1), $\alpha \vdash \neg\neg\alpha$, chaining (Lemma 2.3.3 (c)) gives $\Sigma \vdash \neg\neg\alpha$.
- *Case* $v(\alpha) = \mathrm{F}$. Then $\alpha^v = \neg\alpha$, and $v(\neg\alpha) = \mathrm{T}$, so $\varphi^v = \neg\alpha = \alpha^v$. The hypothesis is already the claim.

Conjunction. Let $\varphi = \alpha \wedge \beta$, and assume $\Sigma \vdash \alpha^v$ and $\Sigma \vdash \beta^v$. By Definition 4.4.2, $v(\alpha \wedge \beta) = \mathrm{T}$ exactly when $v(\alpha) = v(\beta) = \mathrm{T}$.

- *Case* $v(\alpha) = v(\beta) = \mathrm{T}$. Then $\varphi^v = \alpha \wedge \beta$, $\alpha^v = \alpha$, $\beta^v = \beta$. Chaining $\Sigma \vdash \alpha$ and $\Sigma \vdash \beta$ with ($\wedge$1), $\alpha, \beta \vdash \alpha \wedge \beta$, gives $\Sigma \vdash \alpha \wedge \beta$.
- *Case* $v(\alpha) = \mathrm{F}$. Then $\varphi^v = \neg(\alpha \wedge \beta)$ and $\alpha^v = \neg\alpha$. Chaining $\Sigma \vdash \neg\alpha$ with ($\wedge$2), $\neg\alpha \vdash \neg(\alpha \wedge \beta)$, gives $\Sigma \vdash \neg(\alpha \wedge \beta)$. The hypothesis for β is not used.
- *Case* $v(\alpha) = \mathrm{T}$, $v(\beta) = \mathrm{F}$. Then $\varphi^v = \neg(\alpha \wedge \beta)$ and $\beta^v = \neg\beta$, and chaining $\Sigma \vdash \neg\beta$ with ($\wedge$3) gives $\Sigma \vdash \neg(\alpha \wedge \beta)$.

Implication. Let $\varphi = \alpha \to \beta$, and assume $\Sigma \vdash \alpha^v$ and $\Sigma \vdash \beta^v$. By Definition 4.4.2, $v(\alpha \to \beta) = \mathrm{F}$ exactly when $v(\alpha) = \mathrm{T}$ and $v(\beta) = \mathrm{F}$.

- *Case* $v(\alpha) = \mathrm{F}$. Then $\varphi^v = \alpha \to \beta$ and $\alpha^v = \neg\alpha$. Chaining $\Sigma \vdash \neg\alpha$ with $(\to 1)$, $\neg\alpha \vdash \alpha \to \beta$, gives $\Sigma \vdash \alpha \to \beta$.
- *Case* $v(\alpha) = \mathrm{T}$, $v(\beta) = \mathrm{T}$. Then $\varphi^v = \alpha \to \beta$ and $\beta^v = \beta$. Chaining $\Sigma \vdash \beta$ with $(\to 2)$, $\beta \vdash \alpha \to \beta$, gives $\Sigma \vdash \alpha \to \beta$.
- *Case* $v(\alpha) = \mathrm{T}$, $v(\beta) = \mathrm{F}$. Then $\varphi^v = \neg(\alpha \to \beta)$, $\alpha^v = \alpha$ and $\beta^v = \neg\beta$. Chaining $\Sigma \vdash \alpha$ and $\Sigma \vdash \neg\beta$ with $(\to 3)$, $\alpha, \neg\beta \vdash \neg(\alpha \to \beta)$, gives $\Sigma \vdash \neg(\alpha \to \beta)$.

Disjunction and biconditional. These go as for implication. For each pair of values of α and β, the table in the proof of Lemma 4.8.1 names a sequent whose premises are among α^v, β^v and whose conclusion is φ^v: for $\vee$, $(\vee 1)$ if $v(\alpha) = \mathrm{T}$, $(\vee 2)$ if $v(\alpha) = \mathrm{F}$ and $v(\beta) = \mathrm{T}$, $(\vee 3)$ if both are F; for $\leftrightarrow$, $(\leftrightarrow 1)$ to $(\leftrightarrow 4)$ for the value pairs T T, F F, T F, F T. Chaining the induction hypotheses for those premises with that sequent gives $\Sigma \vdash \varphi^v$.

The atoms and $\bot$ have the property, and each formation clause preserves it, so by Theorem 4.3.1 every formula has it. This proves the lemma.

■

Each case used the induction hypothesis only for the parts named in the premises of its sequent: the conjunction case with α false never looked at β. That is the economy of Lemma 4.8.1 at work. The disjunction and biconditional cases are written out in Exercise (C1: The remaining cases).

Here is the lemma run on one formula, with every chained sequent written as a line.

Example. Deriving one row

Let $\varphi = p \to (q \wedge \neg p)$ and let $v(p) = v(q) = \mathrm{T}$. Find φ^v, and write the derivation of $p^v, q^v \vdash \varphi^v$ that the proof of Lemma 4.8.2 produces.

Solution. Compute up the parse tree: $v(\neg p) = \mathrm{F}$, so $v(q \wedge \neg p) = \mathrm{F}$, and since $v(p) = \mathrm{T}$, $v(\varphi) = \mathrm{F}$. So $\varphi^v = \neg(p \to (q \wedge \neg p))$, and the premises are $p^v = p$ and $q^v = q$.

The proof climbs the same tree. At $\neg p$ with $v(p) = \mathrm{T}$, it is the negation case with α true: $(\neg 1)$ gives $\neg\neg p$. At $q \wedge \neg p$, the values are T and F: $(\wedge 3)$ gives $\neg(q \wedge \neg p)$ from $(\neg p)^v = \neg\neg p$. At the root, the values are T and F: $(\to 3)$ gives the goal from p and $\neg(q \wedge \neg p)$. Written as a derivation, with each sequent of Lemma 4.8.1 cited as a derived rule:

1	p	
2	q	
3	$\neg\neg p$	$(\neg 1)$ 1
4	$\neg(q \wedge \neg p)$	$(\wedge 3)$ 3
5	$\neg(p \to (q \wedge \neg p))$	$(\to 3)$ 1, 4

The premise q is never cited. It was used only in computing which case applies: the $\wedge$ case with $\beta = \neg p$ false needs only β's sign, not q's. A derivation need not use every premise (Definition 1.2.3).

Quick check. Let $\varphi = p \to \neg q$ and $v(p) = v(q) = \mathrm{T}$. What sequent does Lemma 4.8.2 provide, and which sequent of Lemma 4.8.1 does the proof use at the root?

Solution. Here $v(\neg q) = \mathrm{F}$, so $v(\varphi) = \mathrm{F}$, and the lemma gives $p, q \vdash \neg(p \to \neg q)$. At the root the values are T and F, so the proof uses $(\to 3)$, $\alpha, \neg\beta \vdash \neg(\alpha \to \beta)$, with $\alpha = p$ and $\beta = \neg q$. Its premise $\neg\beta = \neg\neg q$ comes from the negation case below it: $q \vdash \neg\neg q$, by $(\neg 1)$.

4.8.5 Gluing the rows: splitting on an atom

Lemma 4.8.2 derives a formula from the signed atoms of **one** row. A sequent $\Gamma \vdash \psi$ mentions no row at all. The signed atoms must be removed, and the tool for that is excluded middle.

Lemma 4.8.3. Splitting on an Atom

Let Δ be a list of formulas, p an atom and ψ a formula. If $\Delta, p \vdash \psi$ and $\Delta, \neg p \vdash \psi$, then $\Delta \vdash \psi$.

Idea. This is proof by cases on $p \vee \neg p$. Proposition 2.5.3 gives $p \vee \neg p$ with no premises, and $\vee$E (Definition 1.6.1) needs a box that reaches ψ from p and a box that reaches ψ from $\neg p$. The two given derivations are exactly the contents of those boxes, once their premise p (or $\neg p$) is replaced by the box's assumption.

Before reading on: this is the only place in the completeness proof where a classical rule enters. Where exactly, and why could the $\neg$I *of Chapter 2 not do the job?*

Proof. Let D be a derivation of ψ with premises in Δ, p, and E one with premises in $\Delta, \neg p$. Build a list of lines as follows.

1. As premises, every premise of D or of E that is a member of Δ.
2. The line $p \vee \neg p$, justified by Proposition 2.5.3 (LEM) with p for P; it cites nothing.
3. A box that opens with the assumption p, followed by the lines of D that are not premises, with their boxes. A citation of a premise of D that is in Δ now cites the premise line of step 1; a citation of the premise p now cites the assumption. If the last line of D is itself a premise, the box is shorter: if that premise is p, the box is the assumption alone, which is then its last line; otherwise the box is the assumption followed by ψ, by R, citing the premise line carrying ψ.
4. A second box, built from E in the same way, opening with the assumption $\neg p$.
5. The line ψ, by $\vee$E, citing line 2 of this list and the two boxes.

This is a derivation (Definition 1.4.2). The premises lie in no box and come first, so they are available everywhere. Inside the first box, the lines of D keep their boxes, which now sit inside the new box, and they cite the same lines and boxes as before, or a premise line, or the assumption, which is available at every line of its own box. As in the proof of Lemma 2.3.3, each line is still an instance of its rule. The last line of the first box is ψ, and it is a line of that box itself and not of a box inside it: it was the last line of D, which lies in no box of D (or it is the assumption itself, or the repetition line). The same holds for the second

box. Both boxes and the line $p \vee \neg p$ lie in no box, and neither does the last line, so they are available at it. Hence the last line is an instance of $\vee$E, and the list is a derivation of ψ whose premises are all in Δ. So $\Delta \vdash \psi$.

■

Here is its shape, with Δ standing for the premise lines:

h	Δ	
i	$p \vee \neg p$	LEM
a	$\quad p$	asm, for $\vee$E
	$\quad \vdots$	
b	$\quad \psi$	
c	$\quad \neg p$	asm, for $\vee$E
	$\quad \vdots$	
d	$\quad \psi$	
k	ψ	$\vee$E i, a–b, c–d

The answer to the prompt: the classical rule enters at line i. Proposition 2.5.3 was derived in Chapter 2 with RAA, and nothing else in the proof of this lemma is classical. ¬I could not replace it: ¬I produces a negation, and here we need a disjunction with no premises to split on. Nothing in the proof uses that p is an atom; the lemma holds for any formula in place of p, but atoms are all we shall split on.

4.8.6 The completeness theorem

Now the plan of the opening can be carried out.

Theorem 4.8.4. Completeness

Let Γ be a list of formulas and ψ a formula. If $\Gamma \models \psi$, then $\Gamma \vdash \psi$.

A list is finite, as in Definition 1.2.3. That matters: the proof works with the atoms that occur in Γ and ψ, and it needs them to be finitely many.

Idea. **Step roadmap.** ① List the atoms $p_1, \ldots, p_n$ of Γ and ψ. ② **Claim:** for **every** valuation v, $\Gamma, p_1^v, \ldots, p_n^v \vdash \psi$. If v makes all of Γ true, then ψ is true in that row, and Lemma 4.8.2 derives it. If v makes some premise γ false, the lemma derives $\neg\gamma$, which clashes with the premise γ. ③ Remove the signed atoms one at a time, the last first, by Lemma 4.8.3: an induction on the number of atoms removed. To remove p_j we need derivations from both p_j and $\neg p_j$ with the same other signed atoms: two rows that differ only at p_j. That is why the Claim must hold for every valuation, not just one.

Let us run the method of the roadmap on the tautology of the opening.

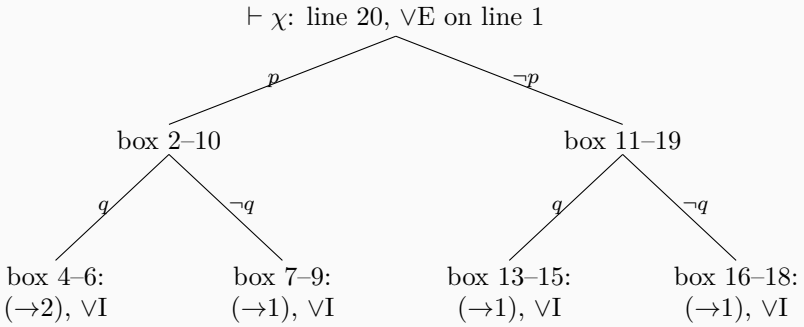
Example. The method on a tautology

Run the roadmap of Theorem 4.8.4 on $\models (p \to q) \vee (q \to p)$, and write the derivation it produces, citing the sequents of Lemma 4.8.1 as derived rules.

Solution. Write χ for $(p \to q) \vee (q \to p)$. The list Γ is empty and the atoms are $p_1 = p$, $p_2 = q$. Every valuation makes χ true (if $v(q) = \mathrm{T}$ the left side is true; otherwise the right side is), so the Claim ② always falls in its first case (the row makes the premises, here none, true), and Lemma 4.8.2 produces four short derivations, one per row:

- $p, q \vdash \chi$: $p \to q$ by $(\to 2)$ from q, then $\vee$I;
- $p, \neg q \vdash \chi$: $q \to p$ by $(\to 1)$ from $\neg q$, then $\vee$I;
- $\neg p, q \vdash \chi$ and $\neg p, \neg q \vdash \chi$: $p \to q$ by $(\to 1)$ from $\neg p$, then $\vee$I.

(In the row $p = \mathrm{F}$, $q = \mathrm{T}$, both $(\to 1)$ and $(\to 2)$ would do; the proof of Lemma 4.8.2 uses $(\to 1)$ whenever α is false.) ③ first splits on $q = p_2$, twice: once with p and once with $\neg p$ among the premises. Then it splits on $p = p_1$, once. Three uses of Lemma 4.8.3 in two rounds. Pasting as in its proof, the last split is the outermost:



Each leaf is one row of the table; each inner node is one use of Lemma 4.8.3.

1	$p \lor \neg p$	LEM
2	$\quad p$	asm, for $\lor$E
3	$\quad q \lor \neg q$	LEM
4	$\quad\quad q$	asm, for $\lor$E
5	$\quad\quad p \to q$	($\to$2) 4
6	$\quad\quad (p \to q) \lor (q \to p)$	$\lor$I 5
7	$\quad\quad \neg q$	asm, for $\lor$E
8	$\quad\quad q \to p$	($\to$1) 7
9	$\quad\quad (p \to q) \lor (q \to p)$	$\lor$I 8
10	$\quad (p \to q) \lor (q \to p)$	$\lor$E 3, 4–6, 7–9
11	$\quad \neg p$	asm, for $\lor$E
12	$\quad q \lor \neg q$	LEM
13	$\quad\quad q$	asm, for $\lor$E
14	$\quad\quad p \to q$	($\to$1) 11
15	$\quad\quad (p \to q) \lor (q \to p)$	$\lor$I 14
16	$\quad\quad \neg q$	asm, for $\lor$E
17	$\quad\quad p \to q$	($\to$1) 11
18	$\quad\quad (p \to q) \lor (q \to p)$	$\lor$I 17
19	$\quad (p \to q) \lor (q \to p)$	$\lor$E 12, 13–15, 16–18
20	$(p \to q) \lor (q \to p)$	$\lor$E 1, 2–10, 11–19

The derivation has 20 lines. The one found by search in Exercise (C2: A disjunction from nothing) has 13, and splits only once. Here the split on q inside the box of $\neg p$ (lines 12–19) is wasted: both of its boxes do the same thing, because in the rows with p false the value of q does not matter. The proof is a **method** that always works, not a recipe for short derivations; to find a short one, search as in Section 3.1.

Quick check. In the tree, which two leaves do the same thing, and why?

Solution. The boxes 13–15 and 16–18. Both write $p \to q$ by ($\to$1) from $\neg p$ on line 11, because in a row with p false, $p \to q$ is true whatever q is.

Proof. **Step 1.** The list Γ is finite and each formula has finitely many symbols, so only finitely many atoms occur in Γ or ψ. List them without repetition as $p_1, \ldots, p_n$, where $n \geq 0$.

Step 2.

Claim. For every valuation v, $\Gamma, p_1^v, \ldots, p_n^v \vdash \psi$.

Proof of Claim. Let v be a valuation, and write Σ for $p_1^v, \ldots, p_n^v$.

Case 1: $v(\gamma) = \mathrm{T}$ for every member γ of Γ. Since $\Gamma \models \psi$, by Definition 4.5.3 $v(\psi) = \mathrm{T}$, so $\psi^v = \psi$. Every atom of ψ is among $p_1, \ldots, p_n$, so Lemma 4.8.2 gives $\Sigma \vdash \psi$, and weakening (Lemma 2.3.3 (a)) gives $\Gamma, \Sigma \vdash \psi$.

Case 2: $v(\gamma) = \mathrm{F}$ for some member γ of Γ. Then $\gamma^v = \neg\gamma$. Every atom of γ is among $p_1, \ldots, p_n$, so Lemma 4.8.2 gives $\Sigma \vdash \neg\gamma$, and by weakening $\Gamma, \Sigma \vdash \neg\gamma$. Also $\Gamma, \Sigma \vdash \gamma$, by the one-line derivation whose only line is the premise γ. By Proposition 2.1.1, $\gamma, \neg\gamma \vdash \psi$. Chaining (Lemma 2.3.3 (c)) gives $\Gamma, \Sigma \vdash \psi$.

■

Step 3. For each integer j with $0 \le j \le n$, let $S(j)$ be the statement

$$\text{for every valuation } v, \quad \Gamma, p_1^v, \ldots, p_j^v \vdash \psi,$$

where for $j = 0$ the list of signed atoms is empty. The Claim is $S(n)$.

Claim. If $1 \le j \le n$ and $S(j)$ holds, then $S(j-1)$ holds.

Proof of Claim. Let v be a valuation. Let v^+ be the valuation that agrees with v at every atom except that $v^+(p_j) = \mathrm{T}$, and v^- the one that agrees with v except that $v^-(p_j) = \mathrm{F}$; by Definition 4.4.1, a valuation may assign any value to each atom, so both exist. The atoms $p_1, \ldots, p_{j-1}$ are different from p_j, so v^+ and v^- agree with v on them, and $p_i^{v^+} = p_i^{v^-} = p_i^v$ for $i < j$. Also $p_j^{v^+} = p_j$ and $p_j^{v^-} = \neg p_j$. Applying $S(j)$ to v^+ and to v^- gives, with Δ standing for the list $\Gamma, p_1^v, \ldots, p_{j-1}^v$,

$$\Delta, p_j \vdash \psi \quad \text{and} \quad \Delta, \neg p_j \vdash \psi.$$

By Lemma 4.8.3, $\Delta \vdash \psi$. Since v was arbitrary, this is $S(j-1)$.

■

Step 2 is $S(n)$. Applying the Claim with $j = n$, then $j = n-1$, and so on down to $j = 1$, removes the signed atoms one at a time and reaches $S(0)$. This descent is an induction on the number of atoms removed; the fold after the proof writes it out.

Step 4. $S(0)$ says that $\Gamma \vdash \psi$, for every valuation v; there is at least one valuation (for instance the one making every atom true), so $\Gamma \vdash \psi$. This proves the theorem.

■

Second reading · *The descent, as an induction*

It remains to go down from $S(n)$ to $S(0)$, which takes n applications of the Claim. We count the applications made: for $m \in \mathbb{N}$, let $T(m)$ say "if $m \le n$, then $S(n-m)$ holds". $T(0)$ says that $S(n)$ holds, which is Step 2. Suppose $T(m)$ holds, and let $m+1 \le n$. Then $m \le n$, so $S(n-m)$ holds by $T(m)$; and $j := n - m$ satisfies $1 \le j \le n$, because $0 \le m \le n-1$. The Claim gives $S(j-1) = S(n-(m+1))$, which is $T(m+1)$. By Axiom 4.2.1, $T(m)$ holds for every $m \in \mathbb{N}$,

and $T(n)$ is $S(0)$.

If $n = 0$, the formulas of Γ and ψ are built from $\bot$ alone, and Step 3 does nothing: $S(0)$ is the Claim itself.

Warning. Completeness promises a derivation, not a short one. For a sequent with n atoms, the proof builds 2^n row derivations and glues them with $2^n - 1$ splittings: with $n = 10$ atoms, that is 1024 row derivations and 1023 uses of LEM. The theorem answers "is there a derivation?", not "how short is the shortest?".

With soundness, completeness closes the circle.

Corollary 4.8.5. Derivable Means Valid

For every list of formulas Γ and every formula ψ,

$$\Gamma \vdash \psi \iff \Gamma \models \psi.$$

In particular, whether $\Gamma \vdash \psi$ holds can be decided by checking the 2^n rows of a truth table, where n is the number of atoms occurring in Γ and ψ.

Proof. ($\Rightarrow$) is Theorem 4.7.2 and ($\Leftarrow$) is Theorem 4.8.4. For the second sentence, let $p_1, \ldots, p_n$ be the atoms occurring in Γ and ψ. By Lemma 4.4.1, the values of the members of Γ and of ψ under a valuation depend only on its values at $p_1, \ldots, p_n$, so $\Gamma \models \psi$ holds exactly when it holds in each of the 2^n rows of the joint truth table. That is a finite check, and by the first sentence it decides $\Gamma \vdash \psi$. ■

So a question that looks as if it needs a search through infinitely many possible derivations can be settled by a finite table. The search for a derivation may still be the quicker way to a **yes**, and one row is the quickest way to a **no** (Corollary 4.7.3), but neither is needed.

Consistent and satisfiable coincide. Section 4.7 showed that a satisfiable list is consistent. Completeness gives the converse. If Δ is unsatisfiable, no row makes Δ true, so $\Delta \models \bot$ (there is no row to check), and $\Delta \vdash \bot$ by Theorem 4.8.4: Δ is inconsistent. So a finite list is consistent exactly when it is satisfiable. The word of Section 3.2 and the word of Section 4.5 pick out the same lists, though one speaks of derivations and the other of rows; Section 3.2's task "derive $\bot$" and Section 4.6's search on a list always give the same answer.

↻ **Second reading** · *Why two definitions of 'follows' are enough*

Before Chapter 1, "ψ follows from Γ" had only an informal meaning: whatever a careful mathematician accepts as a valid inference, by any means. That is not a definition, but two things about it are clear for propositional forms. If $\Gamma \vdash \psi$, then ψ follows informally: each rule is an inference every mathematician accepts, and a chain of accepted inferences is accepted. If

ψ follows informally from Γ by form alone, then $\Gamma \models \psi$: an inference valid by its form cannot have a row, which is one way of reading the atoms, making the premises true and the conclusion false. So the informal notion is squeezed between $\vdash$ and $\models$, and completeness makes the two ends meet: for propositional forms, the informal notion coincides with both. An argument of this shape is due to Georg Kreisel.

Second reading · *Infinitely many premises*

We defined $\Gamma \vdash \psi$ only for a list, that is, for finitely many premises. One may allow an infinite collection Γ of premises and say $\Gamma \vdash \psi$ when ψ is derivable from finitely many members of Γ (a derivation has finitely many lines, so it can only use finitely many). The equivalence with $\Gamma \models \psi$ still holds for infinite Γ (completeness for infinitely many premises); together with the finiteness of derivations it gives the compactness theorem of propositional logic. The optional Section 4.10 proves both; nothing in the core depends on them. Our proof does not reach infinite Γ: Step 1 needs finitely many atoms.

4.8.7 Two columns that agree

Section 4.5 set the two meanings of "follows" side by side. Here is the full table, now that both theorems are proved: they agree row by row (Corollary 4.8.5 and the paragraph on consistency above).

Notion	In terms of derivations	In terms of rows
follows	$\Gamma \vdash \psi$ (Definition 1.2.3, Section 1.2)	$\Gamma \models \psi$ (Definition 4.5.3, Section 4.5)
true by form alone	a theorem, $\vdash \varphi$	a tautology, $\models \varphi$ (Definition 4.5.1)
say the same	provably equivalent (Definition 2.4.2, Section 2.4)	semantically equivalent (Definition 4.5.4)
cannot hold together	inconsistent, $\Gamma \vdash \bot$ (Definition 3.2.1, Section 3.2)	unsatisfiable (Definition 4.5.2)
shown to hold by	one derivation	every row, or a closed search tree (Proposition 4.6.1)
shown to fail by	one refuting row (Corollary 4.7.3)	one refuting row
more premises, cut	Lemma 2.3.3 (Section 2.3): chaining and cut are why a derived rule may be cited	Section 4.5's structural properties: one line about rows each
tied together by	soundness, $\vdash \Rightarrow \models$ (Theorem 4.7.2)	completeness, $\models \Rightarrow \vdash$ (Theorem 4.8.4)

4.8.8 Two roads to every answer

We now have two independent ways to settle any question "is $\Gamma \vdash \psi$?". One is to **search** for a derivation, as in Section 3.1. The other is to **check a table**, and Corollary 4.8.5 says the two always agree. Here are two of Chapter 2's classical laws settled by the table first.

De Morgan's law (d). Is $\neg(p \wedge q) \models \neg p \vee \neg q$? The four rows:

p	q	$\neg(p \wedge q)$	$\neg p \vee \neg q$
T	T	F	F
T	F	T	T
F	T	T	T
F	F	T	T

In every row where the premise is T, so is the conclusion. So $\neg(p \wedge q) \models \neg p \vee \neg q$, and by Theorem 4.8.4, $\neg(p \wedge q) \vdash \neg p \vee \neg q$.

Peirce's law. Is $(p \to q) \to p \models p$? Work backwards from F (Section 4.4). A refuting row would need p false. Then $p \to q$ is true, so $(p \to q) \to p$ is "true implies false", which is false. So no row makes the premise true and the conclusion false: $(p \to q) \to p \models p$, and by Theorem 4.8.4, $(p \to q) \to p \vdash p$.

Both answers came before anyone found a derivation. Chapter 2 found both by hand (Proposition 2.5.5 (d), and Proposition 2.5.4, whose $\vdash ((P \to Q) \to P) \to P$ gives the sequent by $\to$E), and the derivations there use RAA. Completeness explains why a search had to succeed; it does not say how.

Second reading · *What the proof used*

A common alternative route first turns $\Gamma \models \psi$ into a single tautology, $\models \gamma_1 \to (\gamma_2 \to \cdots \to (\gamma_k \to \psi))$, proves completeness for tautologies, and then unpacks the implications with $\to$E. Keeping Γ as premises, as Step 2 does, costs one extra case (a false premise) and saves that detour.

Section 4.5 remarked that the tables do not see which of Chapter 2's laws needed RAA. Now we can say why they cannot. A table decides $\Gamma \models \psi$, and by Corollary 4.8.5 that is the same as derivability with **all** the rules, RAA included. A table has no column recording which rules a derivation uses. Whether a sequent can be derived without RAA is a question about a shorter list of rules, and for that list the completeness theorem is false, as the next paragraph explains. So no truth-table check can tell the laws that need RAA from those that do not.

Which rules did the proof of completeness itself use? Lemma 2.3.3 only pastes derivations. Lemma 4.8.1 uses no RAA, as its proof checks. Lemma 4.8.2 and the Claim chain those sequents, together with Proposition 2.1.1, which Chapter 2 derived before RAA appeared. The one classical step is the line $p \vee \neg p$ of Lemma 4.8.3. So the theorem is really about the full list of rules, RAA included. Without RAA the list is sound but not complete, as the optional Theorem 4.7.4 shows for $\vdash p \vee \neg p$ and $\neg\neg p \vdash p$; the proof above breaks exactly at the splitting.

Warning. Completeness does not say that a tautology has a derivation **without** RAA. It says there is a derivation from the rules of Chapters 1 and 2, all of them. The formula $(p \to q) \vee (q \to p)$ is a tautology, and every derivation found for it so far uses LEM or RAA. Whether a particular tautology can do without them is a different question, and a table cannot answer it.

4.8.9 A truth-table step: TAUT

Completeness has a practical consequence. Any step from earlier lines to a new one that a truth table confirms can be filled in by a derivation, so the step may be written as one line and the filling left to the theorem.

Definition 4.8.1. The Step TAUT

In a derivation of formulas of Definition 4.1.1, a line ψ may be justified by **TAUT**, citing lines $i_1, \dots, i_m$ $(m \geq 0)$ available at it, if $\varphi_{i_1}, \dots, \varphi_{i_m} \models \psi$: every row that makes the cited lines true makes ψ true. For $m = 0$, ψ must be a tautology.

A TAUT line abbreviates a derivation in the rules of Chapters 1 and 2: by Corollary 4.8.5, $\varphi_{i_1}, \dots, \varphi_{i_m} \vdash \psi$, and that derivation is pasted in as a derived rule's is (Definition 2.3.1; Lemma 2.3.3 for a line in no box). So TAUT derives nothing new, and Theorem 4.7.2 holds for derivations that use it. A TAUT line is checked by a table, not found by a search; its justification reads "TAUT" with the cited lines.

Convention. TAUT is a license, not a style. It is used only for a step that a written proof would pass over without comment, such as a rearrangement of connectives ("so if p, then q"), never to hide the step an argument is about.

Example. One line for twelve

Derive $\neg(p \to q) \vdash p \wedge \neg q$ with TAUT, and say what the step abbreviates.

Solution.

1	$\neg(p \to q)$	
2	$p \wedge \neg q$	TAUT 1

The check: the premise is true only in the row $p = \mathrm{T}$, $q = \mathrm{F}$, since an implication is false only there, and $p \wedge \neg q$ is true in that row. Written out with the rules of Chapter 2, the derivation takes twelve lines (fold below). It is a step a written proof passes over ("an implication fails only with a true hypothesis and a false conclusion"), so the convention allows it.

↻ **Second reading** · *The step written out*

An RAA box on $\neg p$, in which Proposition 2.1.2 (VI) gives $p \to q$, clashing with line 1, gives p. A $\neg$I box on q, in which an inner box gives $p \to q$ again, gives $\neg q$. Then $\wedge$I. The proof

that a TAUT line can always be written out builds a longer derivation still, by the method of Theorem 4.8.4; this one was found by search.

1	$\neg(p \to q)$	
2	$\quad \neg p$	asm, for RAA
3	$\quad p \to q$	VI 2
4	$\quad \bot$	¬E 3, 1
5	p	RAA 2–4
6	$\quad q$	asm, for ¬I
7	$\quad\quad p$	asm, for →I
8	$\quad\quad q$	R 6
9	$\quad p \to q$	→I 7–8
10	$\quad \bot$	¬E 9, 1
11	$\neg q$	¬I 6–10
12	$p \wedge \neg q$	∧I 5, 11

Quick check. May line 3 be justified by TAUT: 1 $p \vee q$; 2 $p \to r$; 3 $r \vee q$, TAUT 1, 2?

Solution. Yes. In a row making lines 1 and 2 true, either p is true, and then r is, or p is false, and then q is; either way $r \vee q$ is true.

Warning (Not allowed: TAUT). The first three steps break a clause of Definition 4.8.1, named in the tag; the fourth is legal and breaks the convention.

- "$q \to p$, TAUT 1", with line 1 $p \to q$. The row $p = \mathrm{F}$, $q = \mathrm{T}$ makes line 1 true and the new line false. *Clause "every row that makes the cited lines true".*
- Citing by TAUT a line inside a box that has closed. *Clause "available"*: the scope rule of Definition 1.4.2.
- TAUT on a line that is not a formula of Definition 4.1.1, such as "for every x, $x^2 \geq 0$". *Clause "formulas of Definition 4.1.1".*
- "$\neg p$, TAUT 1, 2", with lines $p \to q$ and $\neg q$, in a proof whose point is that contrapositive step. *Convention*: legal by the definition, but it hides the step the argument is about; write MT (Proposition 2.3.1) or the derivation.

Remark (Looking ahead). Chapter 6 extends TAUT to formulas with quantifiers, reading each quantified part as a letter; Chapter 8 names the steps a written proof passes over without comment.

4.8.10 Exercises

4.8.10.1 A. Check your understanding

Exercise. A1

These questions are about Lemma 4.8.2 and Theorem 4.8.4.

(a) In the conjunction case of the proof of Lemma 4.8.2, suppose $v(\alpha) = \mathrm{F}$ and $v(\beta) = \mathrm{T}$. Which sequent of Lemma 4.8.1 is used, and which induction hypotheses?

(b) Where is RAA used in the proof of Theorem 4.8.4, including inside the results it cites?

(c) State the idea of the proof of Theorem 4.8.4 in two sentences.

(d) Run Step 2 of the proof for $\Gamma = p$, $\psi = p \vee q$ and the valuation with $v(p) = \mathrm{F}$, $v(q) = \mathrm{T}$: which case applies, and which sequent does it prove?

(e) Step 1 lists the atoms **without repetition**. Where in the proof is that used?

(f) Which step of the proof fails if Γ is allowed to be an infinite collection of formulas, each with its own atoms?

Solution. *(a)* $(\wedge 2)$, $\neg\alpha \vdash \neg(\alpha \wedge \beta)$, and only the hypothesis for α, $\Sigma \vdash \neg\alpha$. The hypothesis for β is not needed, because $\alpha \wedge \beta$ is false once α is.

(b) Only inside Lemma 4.8.3, at the line $p \vee \neg p$: Proposition 2.5.3 was derived with RAA. The sequents of Lemma 4.8.1 are derived without it, Lemma 2.3.3 only pastes derivations, and Proposition 2.1.1 comes before RAA in Chapter 2.

(c) Each row of the truth table gives a derivation of the conclusion from the premises and that row's signed atoms: by Lemma 4.8.2 if the row makes every premise true, and from a clash with a false premise otherwise. Two rows that differ at one atom are merged by excluded middle and proof by cases, and after all the atoms are merged the signed atoms are gone.

(d) The valuation makes the premise p false, so Case 2 applies with $\gamma = p$. Lemma 4.8.2 gives $\neg p, q \vdash \neg p$ (the atom case), the premise gives p, and Proposition 2.1.1 gives $p \vee q$. The sequent proved is $p, \neg p, q \vdash p \vee q$.

(e) In the Claim of Step 3: v^+ and v^- change only the value at p_j, and they agree with v on $p_1, \ldots, p_{j-1}$ only because those atoms differ from p_j. If p_j also appeared as some p_i with $i < j$, then $p_i^{v^+} = p_j$ and $p_i^{v^-} = \neg p_j$ would differ, and the two sequents would not share the list Δ that Lemma 4.8.3 needs.

(f) Step 1. Infinitely many atoms may occur, so there is no finite list $p_1, \ldots, p_n$: the Claim would need infinitely many signed atoms as premises, which is not a list, and Step 3's descent from $S(n)$ has no n to start from.

Exercise. A2: Drill

The parts run from signed formulas and the steps of the proof, through consistency and the two columns, to judging TAUT lines.

(a) $v(p) = \mathrm{F}$, $v(q) = \mathrm{T}$. Write p^v, q^v, $(p \to q)^v$, $(q \to p)^v$.

(b) $\varphi = \neg p \wedge q$, $v(p) = v(q) = \mathrm{T}$. Which sequent does Lemma 4.8.2 give?

(c) In (b), which tagged sequent does the proof use at the root, and where does its premise come from?

(d) A sequent has three atoms. How many row derivations and how many uses of Lemma 4.8.3 does the method build?

(e) In the step from $S(j)$ to $S(j-1)$, which signed atom disappears?

(f) $\Gamma = q$, $\psi = p \vee q$, $v(p) = v(q) = \mathrm{F}$. Which case of the Claim of Step 2 applies?

(g) Is this TAUT line legal: 2 $p \vee q$; 3 $\neg p$; 5 q, TAUT 2, 3?

(h) Is this one legal: 1 $p \to q$; 4 $q \to p$, TAUT 1?

(i) Is the list $p \vee q$, $\neg p$, $\neg q$ consistent?

(j) In the table "Two columns that agree", what is the twin of "a theorem"?

(k) May $\neg(p \wedge \neg p)$ be written by TAUT citing no line?

(l) A list is satisfiable. What does completeness add to what soundness already says about it?

Solution.

(a) $\neg p$, q, $p \to q$, $\neg(q \to p)$.

(b) $p, q \vdash \neg(\neg p \wedge q)$.

(c) ($\wedge$2) with $\alpha = \neg p$: $\neg\neg p \vdash \neg(\neg p \wedge q)$. The premise $\neg\neg p$ comes from p by ($\neg$1).

(d) 8 and 7.

(e) p_j^v.

(f) Case 2, with $\gamma = q$, which v makes false.

(g) Yes: every row making $p \vee q$ and $\neg p$ true makes q true.

(h) No: the row $p = \mathrm{F}$, $q = \mathrm{T}$ makes line 1 true and line 4 false.

(i) No. It is unsatisfiable, so inconsistent ("Consistent and satisfiable coincide"); directly, $\vee$E on $p \vee q$ with $\neg$E in each box gives $\bot$.

(j) A tautology.

(k) Yes: it is a tautology ($m = 0$).

(l) Nothing about this list: soundness already makes it consistent. Completeness speaks of unsatisfiable lists: they are inconsistent.

4.8.10.2 B. Practice

Exercise. B1: Three rows from the primitive rules

The table in the proof of Lemma 4.8.1 obtains ($\wedge$2), ($\wedge$3) and ($\vee$3) from De Morgan's laws. Derive

$$\neg\alpha \vdash \neg(\alpha \wedge \beta), \qquad \neg\beta \vdash \neg(\alpha \wedge \beta), \qquad \neg\alpha, \neg\beta \vdash \neg(\alpha \vee \beta)$$

directly from the primitive rules, and say which of your derivations use RAA.

Solution. Each goal is a negation, so each derivation opens a $\neg$I box on the unnegated formula and hunts for a clash.

1	$\neg\alpha$		1	$\neg\beta$	
2	$\quad\alpha \wedge \beta$	asm, for $\neg$I	2	$\quad\alpha \wedge \beta$	asm, for $\neg$I
3	$\quad\alpha$	$\wedge$E 2	3	$\quad\beta$	$\wedge$E 2
4	$\quad\bot$	$\neg$E 3, 1	4	$\quad\bot$	$\neg$E 3, 1
5	$\neg(\alpha \wedge \beta)$	$\neg$I 2–4	5	$\neg(\alpha \wedge \beta)$	$\neg$I 2–4

For the disjunction, the assumption $\alpha \vee \beta$ is used by $\vee$E with the goal $\bot$, and each case clashes with one premise.

1	$\neg\alpha$	
2	$\neg\beta$	
3	$\quad\alpha \vee \beta$	asm, for $\neg$I
4	$\quad\quad\alpha$	asm, for $\vee$E
5	$\quad\quad\bot$	$\neg$E 4, 1
6	$\quad\quad\beta$	asm, for $\vee$E
7	$\quad\quad\bot$	$\neg$E 6, 2
8	$\quad\bot$	$\vee$E 3, 4–5, 6–7
9	$\neg(\alpha \vee \beta)$	$\neg$I 3–8

The three derivations have 5, 5 and 9 lines, and none uses RAA: every box closes by $\neg$I or $\vee$E.

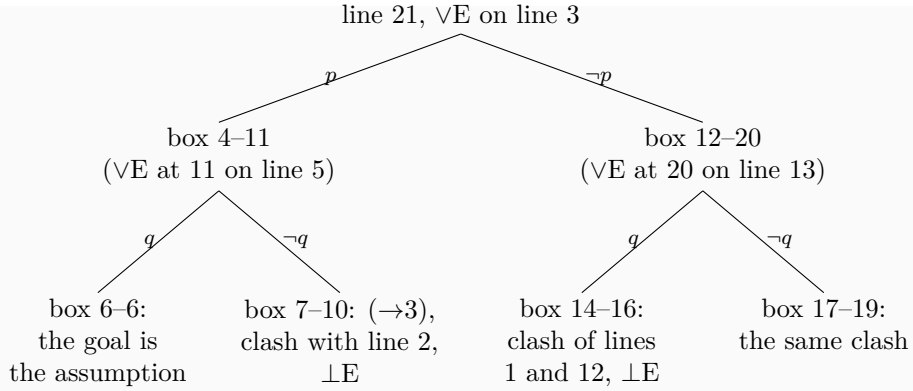
Exercise. B2: The method on modus ponens

Apply the method of Theorem 4.8.4 to $p,\ p \to q \models q$: list the four rows and the case of Step 2 that each falls under, and write the derivation that the proof produces, citing LEM and the sequents of Lemma 4.8.1 as single lines. (It is long; that is expected, and the usual length of an exercise answer does not apply.) Compare it with the derivation by $\to$E.

Solution. The atoms are $p_1 = p$ and $p_2 = q$. The rows:

- $p = \mathrm{T}$, $q = \mathrm{T}$: both premises true, Case 1; $q^v = q$ is itself a premise.
- $p = \mathrm{T}$, $q = \mathrm{F}$: the premise $p \to q$ is false, Case 2 with $\gamma = p \to q$; Lemma 4.8.2 gives $p, \neg q \vdash \neg(p \to q)$ by $(\to 3)$.
- $p = \mathrm{F}$, q either value: the premise p is false, Case 2 with $\gamma = p$; the lemma gives $\neg p$ (an atom case).

In Case 2, the clash is $\neg$E followed by $\bot$E, which is the derivation of Proposition 2.1.1 written out. Splitting on q inside each case of a split on p, the derivation has this shape:



Written out:

1	p	
2	$p \to q$	
3	$p \vee \neg p$	LEM
4	$\quad p$	asm, for ∨E
5	$\quad q \vee \neg q$	LEM
6	$\quad\quad q$	asm, for ∨E
7	$\quad\quad \neg q$	asm, for ∨E
8	$\quad\quad \neg(p \to q)$	(→3) 4, 7
9	$\quad\quad \bot$	¬E 2, 8
10	$\quad\quad q$	⊥E 9
11	$\quad q$	∨E 5, 6–6, 7–10
12	$\quad \neg p$	asm, for ∨E
13	$\quad q \vee \neg q$	LEM
14	$\quad\quad q$	asm, for ∨E
15	$\quad\quad \bot$	¬E 1, 12
16	$\quad\quad q$	⊥E 15
17	$\quad\quad \neg q$	asm, for ∨E
18	$\quad\quad \bot$	¬E 1, 12
19	$\quad\quad q$	⊥E 18
20	$\quad q$	∨E 13, 14–16, 17–19
21	q	∨E 3, 4–11, 12–20

The box at line 6 has one line: in the row $p = q = \mathrm{T}$ the goal q is the signed atom itself, so, as in the proof of Lemma 4.8.3, the box is its assumption alone. The result has 21 lines. The derivation by →E has 3: the two premises and q by →E citing lines 2 and 1. The method is uniform, not short: it does the same thing for every valid sequent, and pays for that with a split on every atom.

Exercise. B3: Table first, then a derivation

Decide by a truth table whether $\neg(p \leftrightarrow q) \models \neg p \leftrightarrow q$. If it holds, find a derivation of $\neg(p \leftrightarrow q) \vdash \neg p \leftrightarrow q$ by search, where you may cite the sequents of Lemma 4.8.1 as derived rules. Which rule of your derivation is classical?

Solution. The table:

p	q	$\neg(p \leftrightarrow q)$	$\neg p \leftrightarrow q$
T	T	F	F
T	F	T	T
F	T	T	T
F	F	F	F

The premise is true in rows 2 and 3, and so is the conclusion. So the entailment holds, and by Theorem 4.8.4 a derivation exists.

The search: the goal is a biconditional, so two boxes, one from $\neg p$ to q and one from q to $\neg p$. In the second, the goal $\neg p$ is a negation: a box on p, and p, q give $p \leftrightarrow q$ by ($\leftrightarrow$1), which clashes with the premise. In the first, the goal q is an atom, the only premise is a negation, and nothing else applies: we are stuck, and assume $\neg q$. Then $\neg p$, $\neg q$ give $p \leftrightarrow q$ by ($\leftrightarrow$2), which clashes with the premise.

1	$\neg(p \leftrightarrow q)$	
2	$\quad \neg p$	asm, for $\leftrightarrow$I
3	$\quad\quad \neg q$	asm, for RAA
4	$\quad\quad p \leftrightarrow q$	($\leftrightarrow$2) 2, 3
5	$\quad\quad \bot$	$\neg$E 4, 1
6	$\quad q$	RAA 3–5
7	$\quad q$	asm, for $\leftrightarrow$I
8	$\quad\quad p$	asm, for $\neg$I
9	$\quad\quad p \leftrightarrow q$	($\leftrightarrow$1) 8, 7
10	$\quad\quad \bot$	$\neg$E 9, 1
11	$\quad \neg p$	$\neg$I 8–10
12	$\neg p \leftrightarrow q$	$\leftrightarrow$I 2–6, 7–11

The classical rule is the RAA at line 6. (Pasting in the derivations of the two cited sequents gives a derivation in the primitive rules of 22 lines.) Whether some other derivation could avoid RAA is a question the table cannot answer. In fact none can; we state that without proof, as Chapter 2 did for similar sequents, and nothing depends on it.

Exercise. B4: When TAUT is welcome

(a) Derive $p \to (q \wedge r),\ \neg r \vdash \neg p$ twice: once with a single TAUT line, once with primitive rules.

(b) By the convention of this section, which is the better proof, and why?

(c) In a longer derivation, line 7 holds $\neg(p \wedge \neg q)$ and line 8 holds $p \to q$. Justify line 8 by TAUT, and say why the convention allows it.

Solution.

(a) With TAUT, line 3 is justified because a row making both premises true has r false, so $q \wedge r$ false, so p false. With primitive rules, a $\neg$I box on p:

1	$p \to (q \wedge r)$	
2	$\neg r$	
3	$\neg p$	TAUT 1, 2

1	$p \to (q \wedge r)$	
2	$\neg r$	
3	$\quad p$	asm, for $\neg$I
4	$\quad q \wedge r$	$\to$E 1, 3
5	$\quad r$	$\wedge$E 4
6	$\quad \bot$	$\neg$E 5, 2
7	$\neg p$	$\neg$I 3–6

(b) The primitive one. The TAUT line hides the only idea of the argument, modus tollens through the conjunction, which is exactly what the convention forbids.

(c) 8 $p \to q$, TAUT 7. The two formulas are equivalent (each is false exactly at $p = \mathrm{T}$, $q = \mathrm{F}$), and the step is a rearrangement a written proof makes without comment.

Marking scheme. (10 points)

- 3: (a) the TAUT derivation, with the row check.
- 2: (a) the primitive derivation.
- 3: (b) the primitive one, because the TAUT line hides the argument's only idea.
- 2: (c) the TAUT line, with why the convention allows it.

Exercise. B5: Consistent or not

Decide whether each list is consistent. Justify "consistent" by a row and the result that makes it a proof; justify "inconsistent" by a derivation of $\bot$, or by a search and the result that makes it a proof.

(a) $p \to q,\ q \to r,\ p,\ \neg r$.

(b) $p \vee q,\ \neg p \vee \neg q,\ p \leftrightarrow q$.

(c) $p \to \neg q,\ q \to \neg r,\ r \to \neg p$.

Solution.

(a) Inconsistent:

1	$p \to q$	
2	$q \to r$	
3	p	
4	$\neg r$	
5	q	$\to$E 1, 3
6	r	$\to$E 2, 5
7	$\bot$	$\neg$E 6, 4

(b) Inconsistent. A search closes: $p \leftrightarrow q$ gives p, q or $\neg p, \neg q$; under the first, $\neg p \vee \neg q$ closes both ways; under the second, $p \vee q$ closes both ways. So the list is unsatisfiable (Proposition 4.6.1), hence inconsistent, by Theorem 4.8.4 ("Consistent and satisfiable coincide").

(c) Consistent. The row with p, q, r all false makes all three members true (false hypotheses), so the list is satisfiable, hence consistent (Section 4.7, by Theorem 4.7.2).

Marking scheme. (10 points)

- 3: (a) the verdict (1) and the derivation of $\bot$ (2).
- 4: (b) the verdict (1) and the closed search with the results that make it a proof (3).
- 3: (c) the verdict (1) and the row with the theorem it rests on (2).

4.8.10.3 C. Going deeper

Exercise. C1: The remaining cases

Write out in full the disjunction case and the biconditional case of the proof of Lemma 4.8.2, in the style of the implication case.

Solution. Throughout, Σ is $p_1^v, \ldots, p_n^v$, and the induction hypotheses are $\Sigma \vdash \alpha^v$ and $\Sigma \vdash \beta^v$. "Chaining" is Lemma 2.3.3 (c).

Disjunction, $\varphi = \alpha \vee \beta$. By Definition 4.4.2, $v(\alpha \vee \beta) = \mathrm{F}$ exactly when $v(\alpha) = v(\beta) = \mathrm{F}$.

- *Case* $v(\alpha) = \mathrm{T}$. Then $\varphi^v = \alpha \vee \beta$ and $\alpha^v = \alpha$. Chaining $\Sigma \vdash \alpha$ with $(\vee 1)$, $\alpha \vdash \alpha \vee \beta$, gives $\Sigma \vdash \alpha \vee \beta$.
- *Case* $v(\alpha) = \mathrm{F}$, $v(\beta) = \mathrm{T}$. Then $\varphi^v = \alpha \vee \beta$ and $\beta^v = \beta$. Chaining $\Sigma \vdash \beta$ with $(\vee 2)$ gives $\Sigma \vdash \alpha \vee \beta$.
- *Case* $v(\alpha) = v(\beta) = \mathrm{F}$. Then $\varphi^v = \neg(\alpha \vee \beta)$, $\alpha^v = \neg\alpha$ and $\beta^v = \neg\beta$. Chaining both hypotheses with $(\vee 3)$, $\neg\alpha, \neg\beta \vdash \neg(\alpha \vee \beta)$, gives $\Sigma \vdash \neg(\alpha \vee \beta)$.

Biconditional, $\varphi = \alpha \leftrightarrow \beta$. By Definition 4.4.2, $v(\alpha \leftrightarrow \beta) = \mathrm{T}$ exactly when $v(\alpha) = v(\beta)$. Each case uses both hypotheses.

- *Case T, T.* $\varphi^v = \alpha \leftrightarrow \beta$; chain $\Sigma \vdash \alpha$, $\Sigma \vdash \beta$ with $(\leftrightarrow 1)$.
- *Case F, F.* $\varphi^v = \alpha \leftrightarrow \beta$; chain $\Sigma \vdash \neg\alpha$, $\Sigma \vdash \neg\beta$ with $(\leftrightarrow 2)$.
- *Case T, F.* $\varphi^v = \neg(\alpha \leftrightarrow \beta)$; chain $\Sigma \vdash \alpha$, $\Sigma \vdash \neg\beta$ with $(\leftrightarrow 3)$.
- *Case F, T.* $\varphi^v = \neg(\alpha \leftrightarrow \beta)$; chain $\Sigma \vdash \neg\alpha$, $\Sigma \vdash \beta$ with $(\leftrightarrow 4)$.

The cases cover all four pairs of values, so in each the property holds for φ.

Marking scheme. (10 points)

- 4: the $\vee$ case: the three value cases, each with the right sequent of Lemma 4.8.1 and only the hypotheses its premises need.
- 4: the $\leftrightarrow$ case: the four value cases, each with the right sequent and both hypotheses.
- 1: names chaining (Lemma 2.3.3 (c)) as the step that combines the hypotheses with the sequent.
- 1: says that the cases cover every pair of values.

Fatal errors (0 for the step): using a sequent whose premises are not the signed parts α^v, β^v of that case.

Exercise. C2: Exit check: completeness

Close the section first, and write from memory.

(a) Prove Lemma 4.8.3: if $\Delta, p \vdash \psi$ and $\Delta, \neg p \vdash \psi$, then $\Delta \vdash \psi$.

(b) Prove Step 3 of the proof of Theorem 4.8.4 in full: from the Claim of Step 2, for every valuation v, $\Gamma, p_1^v, \ldots, p_n^v \vdash \psi$, deduce $\Gamma \vdash \psi$. State the statement you prove by induction, and say which principle of induction you use.

Solution (Hint 1). For (a): excluded middle and proof by cases. For (b): remove the signed atoms from the end of the list, one at a time.

Solution (Hint 2). For (b): let $S(j)$ say "for every valuation v, $\Gamma, p_1^v, \ldots, p_j^v \vdash \psi$", and show that $S(j)$ implies $S(j-1)$ for $1 \leq j \leq n$.

Solution. *(a)* Let D derive ψ from premises in Δ, p, and E from premises in $\Delta, \neg p$. Write, as premises, those of D and E that lie in Δ; then $p \vee \neg p$ by LEM (Proposition 2.5.3, a derived rule with no premises); then a box opening with p containing the non-premise lines of D, in which citations of the premise p now cite the assumption and citations of premises in Δ cite the premise lines; then a box opening with $\neg p$, built from E in the same way; then ψ by $\vee$E, citing the line $p \vee \neg p$ and the two boxes. (If D or E ends on a premise, its box is the assumption alone when that premise is the assumption's formula, and otherwise the assumption followed by ψ by R.) Every copied line keeps its boxes, now inside the new box, and cites lines still available to it. The last line of each box is ψ, as a line of that box itself, because the last line of D or E lay in no box. The two boxes and the LEM line lie in no box, so they are available at the final line. So the result is a derivation of ψ from premises in Δ, and $\Delta \vdash \psi$.

(b) For $0 \leq j \leq n$, let $S(j)$ be: for every valuation v, $\Gamma, p_1^v, \ldots, p_j^v \vdash \psi$. The Claim of Step 2 is $S(n)$.

The step. Let $1 \leq j \leq n$ and assume $S(j)$. Let v be any valuation, and let v^+, v^- agree with v except that $v^+(p_j) = \mathrm{T}$ and $v^-(p_j) = \mathrm{F}$. Since the p_i are distinct, both agree with v on $p_1, \ldots, p_{j-1}$. By $S(j)$ for v^+ and for v^-, with $\Delta := \Gamma, p_1^v, \ldots, p_{j-1}^v$: $\Delta, p_j \vdash \psi$ and $\Delta, \neg p_j \vdash \psi$. By (a), $\Delta \vdash \psi$. As v was arbitrary, $S(j-1)$ holds.

The induction. We use Axiom 4.2.1 on $m \in \mathbb{N}$ with the statement $T(m)$: "if $m \leq n$, then $S(n-m)$". $T(0)$ is $S(n)$. If $T(m)$ holds and $m+1 \leq n$, then $S(n-m)$ holds and $j = n-m$ lies between 1 and n, so the step gives $S(n-m-1)$, which is $T(m+1)$. Hence $T(n)$, which is $S(0)$: for every valuation, $\Gamma \vdash \psi$. Valuations exist, so $\Gamma \vdash \psi$.

(Inducting downwards on j, or on the number of atoms removed, is the same argument; either earns full marks.)

Marking scheme. (10 points)

- 3: (a) builds one derivation from the two given ones with premises in Δ (for instance by pasting them into the two boxes of a $\vee$E on $p \vee \neg p$, or by an RAA box on $\neg\psi$), with each premise p or $\neg p$ replaced by an available line, and says why every citation is still available.
- 1: (a) names the classical step (LEM, Proposition 2.5.3, or RAA).
- 3: (b) sets up an induction on the number of atoms removed (or downward on j) with a statement quantified over **every** valuation, and names the induction principle.
- 2: (b) in the step, applies the hypothesis to **both** valuations that differ at p_j, and notes that they agree with v on the other atoms.
- 1: (b) concludes $\Gamma \vdash \psi$ from $S(0)$.

Fatal errors (0 for the step): splitting on an atom using the Claim for a single valuation only; an induction statement that fixes v before the induction, so that the two valuations of the step are not covered by the hypothesis.

Exercise. C3: Exit check: deciding sequents

For each sequent, decide whether it holds. If it does, justify this by a derivation, or by a truth table and the theorem that turns the table into a derivation; if it does not, give a refuting row and say which theorem turns the row into a proof that no derivation exists.

(a) $(p \to q) \vee (p \to r) \vdash p \to (q \vee r)$.
(b) $(p \to q) \to r \vdash p \to r$.
(c) $p \vee (q \wedge r) \vdash (p \vee q) \wedge r$.
(d) $(p \vee q) \wedge \neg p \vdash q \vee r$.

Solution. (a) Derivable. The goal is an implication: a box on p, with the goal $q \vee r$; the premise is a disjunction, used by $\vee$E inside the box; in each case $\to$E with p gives one side, and $\vee$I gives the goal.

1	$(p \to q) \vee (p \to r)$	
2	$\quad p$	asm, for $\to$I
3	$\quad\quad p \to q$	asm, for $\vee$E
4	$\quad\quad q$	$\to$E 3, 2
5	$\quad\quad q \vee r$	$\vee$I 4
6	$\quad\quad p \to r$	asm, for $\vee$E
7	$\quad\quad r$	$\to$E 6, 2
8	$\quad\quad q \vee r$	$\vee$I 7
9	$\quad q \vee r$	$\vee$E 1, 3–5, 6–8
10	$p \to (q \vee r)$	$\to$I 2–9

(b) Not derivable. In the row $p = \mathrm{T}$, $q = \mathrm{F}$, $r = \mathrm{F}$, the premise is true ($p \to q$ is F, so $(p \to q) \to r$ is T) and the conclusion $p \to r$ is F. By Corollary 4.7.3, $(p \to q) \to r \nvdash p \to r$.

(c) Not derivable. In the row $p = \mathrm{T}$, $q = \mathrm{T}$, $r = \mathrm{F}$ (the value of q does not matter), the premise is true because p is, and the conclusion is false because r is. By Corollary 4.7.3, the sequent fails.

(d) Derivable. Split the premise by $\wedge$E, then use $p \vee q$ by $\vee$E: the case p clashes with $\neg p$ and closes by $\bot$E; the case q gives the goal by $\vee$I.

1	$(p \vee q) \wedge \neg p$	
2	$p \vee q$	$\wedge$E 1
3	$\neg p$	$\wedge$E 1
4	$\quad p$	asm, for $\vee$E
5	$\quad \bot$	$\neg$E 4, 3
6	$\quad q \vee r$	$\bot$E 5
7	$\quad q$	asm, for $\vee$E
8	$\quad q \vee r$	$\vee$I 7
9	$q \vee r$	$\vee$E 2, 4–6, 7–8

Neither derivation uses RAA. As the problem allows, (a) and (d) may instead be justified by a table (or backwards from F) showing the entailment, followed by Theorem 4.8.4.

Marking scheme. (10 points; 2.5 per sequent)

- 1 per sequent: the correct verdict.
- 1.5 per sequent: its justification: a derivation in which every line is justified and every citation available, or a table check with Theorem 4.8.4; or a row that makes **every** premise true and the conclusion false, checked, with Corollary 4.7.3 (or Theorem 4.7.2) cited.

Fatal errors (0 for the sequent's justification): a "refuting row" that makes some premise

false; a derivation citing a line that is not available.

Exercise. C4: Challenge: Strike out $\vee$I

Strike $\vee$I from the primitive rules.

(a) Which sequents of Lemma 4.8.1 lose the derivations given in its proof (the table of the fold "The fifteen derivations"), and which step of the proof of Theorem 4.8.4 also uses $\vee$I, through a result of Chapter 2?

(b) Read $\vee$ as $\wedge$: a valuation now makes $\varphi \vee \psi$ true exactly when it makes both parts true; every other clause stays. Show that every remaining primitive rule is sound for this reading.

(c) Deduce that neither $p \vdash p \vee q$ nor $\vdash p \vee \neg p$ is derivable without $\vee$I (and without derived rules derived with it), although both are valid: the shorter list of rules is sound but not complete.

Solution.

(a) ($\vee$1) and ($\vee$2) are $\vee$I itself; ($\wedge$2) and ($\wedge$3) are derived by $\vee$I and then Proposition 2.5.5 (c). And Lemma 4.8.3 uses Proposition 2.5.3, whose derivation in Section 2.5 uses $\vee$I twice; that is Step 3 of the proof, where Lemma 4.8.3 removes the signed atoms.

(b) Only the cases $\vee$I and $\vee$E of the proof of Theorem 4.7.2 use the clause for $\vee$, and $\vee$I is gone. $\vee$E: if a valuation makes Γ_k true, it makes $\varphi \vee \psi$ true, hence, in the new reading, both φ and ψ; by the first box it makes χ true. Every other case is unchanged.

(c) The valuation with p true and q false makes p true and, in the new reading, $p \vee q$ false; every valuation makes $p \vee \neg p$ false in the new reading, since p and $\neg p$ are never both true. By (b), and Step 1 of the proof of Theorem 4.7.2 (a derived rule derived without $\vee$I expands, by that step, into a pasted derivation in the remaining primitive rules), neither sequent has a derivation without $\vee$I and the derived rules that use it, such as Proposition 2.5.3. Both are valid in the ordinary tables.

Marking scheme. (10 points)

- 3: (a), all four sequents and the excluded-middle step.
- 4: (b), with the $\vee$E case argued and the other cases said to be unchanged.
- 3: (c), both valuations and the conclusion, with Step 1's restriction to derived rules derived without $\vee$I.

Fatal errors (0 for the step): in (c), using the ordinary tables.

Exercise. C5: Challenge: From a closed tree to a derivation

Let a search tree (Section 4.6) for $\Gamma, \neg\psi$ have every branch closed. For a node X, the *path list* of X is the list of formulas on the path from the root down to X, X included.

(a) Prove, by induction on the number of nodes below X, that if every branch through X is closed, then $\Delta \vdash \bot$, where Δ is the path list of X. Treat a closed leaf and each rule of the table of search rules, naming the derivable sequents you use.

(b) Deduce $\Gamma \vdash \psi$.

(c) Explain why (a), (b) and Proposition 4.6.1 give a second proof of Theorem 4.8.4, and where it uses RAA or LEM.

Solution (Hint 1). A one-way rule: derive the added formulas from the used one, then chain (Lemma 2.3.3). A two-way rule: join the two derivations of $\bot$.

Solution (Hint 2). For $\varphi \to \psi$, split as $\neg\varphi$ or ψ: with Δ the path list of X, from $\Delta, \neg\varphi \vdash \bot$ get $\Delta \vdash \varphi$ by RAA.

Solution.

(a) We use Theorem 4.2.2 on the number of nodes below X. Write Δ for the path list of X.
A closed leaf. Δ holds φ and $\neg\varphi$ (one $\neg$E line) or $\bot$ (the premise itself): $\Delta \vdash \bot$.
An inner node. Its children were added by the rule of some χ in Δ. Every branch through each child is closed, and each child has fewer nodes below it, so the hypothesis gives $\Delta, \alpha_1 \vdash \bot$ (or $\Delta, \alpha_1, \alpha_2 \vdash \bot$) for the formulas α that each child adds.
One-way rules. It suffices to show $\Delta \vdash \alpha$ for each added α, and then chain (Lemma 2.3.3 (c), after weakening). $\varphi \wedge \psi$: $\wedge$E. $\neg(\varphi \vee \psi) \vdash \neg\varphi$: a $\neg$I box on φ, $\vee$I, $\neg$E; likewise $\neg\psi$. $\neg(\varphi \to \psi) \vdash \varphi$: an RAA box on $\neg\varphi$, in which Proposition 2.1.2 gives $\varphi \to \psi$, then $\neg$E. $\neg(\varphi \to \psi) \vdash \neg\psi$: a $\neg$I box on ψ, in which an inner box gives $\varphi \to \psi$ by R and $\to$I, then $\neg$E. $\neg\neg\varphi \vdash \varphi$: Proposition 2.5.1.
Two-way rules. $\varphi \vee \psi$: $\vee$E on the line $\varphi \vee \psi$, with the two derivations of $\bot$ pasted into the boxes as in the proof of Lemma 4.8.3. $\neg(\varphi \wedge \psi)$: from $\Delta, \neg\varphi \vdash \bot$, $\Delta \vdash \varphi$ by RAA; likewise $\Delta \vdash \psi$; then $\wedge$I, and $\neg$E with $\neg(\varphi \wedge \psi)$. $\varphi \to \psi$: $\Delta \vdash \varphi$ by RAA from the first child; $\to$E gives ψ; cut with $\Delta, \psi \vdash \bot$. $\varphi \leftrightarrow \psi$: from $\Delta, \varphi, \psi \vdash \bot$ and $\Delta, \varphi \vdash \psi$ (by $\leftrightarrow$E), cut gives $\Delta, \varphi \vdash \bot$; from $\Delta, \neg\varphi, \neg\psi \vdash \bot$ and $\Delta, \neg\varphi \vdash \neg\psi$ (a $\neg$I box on ψ: φ by $\leftrightarrow$E, clash), cut gives $\Delta, \neg\varphi \vdash \bot$; join the two by Lemma 4.8.3, which holds for any formula in place of the atom (the remark after its proof). $\neg(\varphi \leftrightarrow \psi)$: $\Delta, \varphi \vdash \neg\psi$ (a $\neg$I box on ψ: $\varphi \leftrightarrow \psi$ by $\leftrightarrow$I with two two-line boxes, each closed by R, then the clash with $\neg(\varphi \leftrightarrow \psi)$), so $\Delta, \varphi \vdash \bot$ by cut with the first child; $\Delta, \neg\varphi \vdash \psi$ (an RAA box on $\neg\psi$: $\varphi \leftrightarrow \psi$ by $\leftrightarrow$I, each box closing by a clash and $\bot$E, then the clash with $\neg(\varphi \leftrightarrow \psi)$), so $\Delta, \neg\varphi \vdash \bot$ by cut with the second child; join by Lemma 4.8.3.

(b) At the root, $\Gamma, \neg\psi \vdash \bot$; an RAA box on $\neg\psi$ holding that derivation gives $\Gamma \vdash \psi$ (Proposition 3.2.2 (b)).

(c) If $\Gamma \models \psi$, a search tree for $\Gamma, \neg\psi$ can be continued until every branch is closed or finished (Proposition 4.6.1 (a)), and it has no open finished branch, since by (c) of that proposition its row would refute $\Gamma \models \psi$. So every branch is closed, and (a), (b) give $\Gamma \vdash \psi$. RAA enters in the cases $\neg(\varphi \to \psi)$, $\neg\neg\varphi$, $\neg(\varphi \wedge \psi)$, $\varphi \to \psi$, $\neg(\varphi \leftrightarrow \psi)$ and at the root; LEM in the two biconditional cases. The derivation follows the tree, so it is often far shorter than the 2^n rows of the first proof.

Marking scheme. (10 points)

- 1: the closed leaf.
- 2: the four one-way cases.

- 5: the five two-way cases (1 each).
- 1: (b).
- 1: (c), with the classical steps located.

Fatal errors (0 for the step): treating a split by assuming one side only.

4.8.10.4 R. Review

Exercise. R1

Find a derivation of $p \vee q,\ \neg(p \wedge r),\ r \vdash q$.

Solution. The method of Section 3.1 (Example (Proof by cases, found by the method) shows it on a proof by cases). The goal q is an atom, so no introduction rule applies; the premise $p \vee q$ is a disjunction, so we use it by $\vee$E, with the goal q in both boxes. The case q is the goal. In the case p, nothing gives q directly, so we look for a clash: p and the premise r build $p \wedge r$ by $\wedge$I, which clashes with $\neg(p \wedge r)$, and $\bot$E closes the impossible case (Example (Manufacturing the other half of a clash)).

1	$p \vee q$	
2	$\neg(p \wedge r)$	
3	r	
4	$\quad p$	asm, for $\vee$E
5	$\quad p \wedge r$	$\wedge$I 4, 3
6	$\quad \bot$	$\neg$E 5, 2
7	$\quad q$	$\bot$E 6
8	$\quad q$	asm, for $\vee$E
9	q	$\vee$E 1, 4–7, 8–8

The second box is its assumption alone, which is already the goal. No RAA is used.

Exercise. R2

Let m and n be integers, and consider the argument: "$m + n$ is even only if m and n are both even or both odd. m is even and n is odd. Therefore $m + n$ is not even." Translate it into a sequent, using atoms e: "$m + n$ is even", a: "m is even", b: "n is even", and decide whether the premises entail the conclusion.

Solution. "P only if Q" is $P \to Q$ (Section 1.1, and Exercise (C1: "Only if" and "unless", used later)). "Both odd" is $\neg a \wedge \neg b$, since an integer is odd exactly when it is not even. So the sequent is

$$e \to ((a \wedge b) \vee (\neg a \wedge \neg b)),\quad a \wedge \neg b \models \neg e.$$

Work backwards from F. A refuting row needs $\neg e$ false, so $e =$ T, and the second premise

true, so $a = \mathrm{T}$, $b = \mathrm{F}$. In that row $a \wedge b$ and $\neg a \wedge \neg b$ are both F, so the first premise is $\mathrm{T} \to \mathrm{F}$, which is F. So no row makes both premises true and the conclusion false: the entailment holds. By Theorem 4.8.4, the sequent is also derivable. (The eight-row table gives the same answer; working backwards found the only candidate row at once.)

Exercise. R3

Is $(p \wedge (q \vee r)) \to ((p \wedge q) \vee r)$ a tautology? Is it derivable without RAA?

Solution. A tautology. Work backwards from F (Section 4.4). The formula is false only if $p \wedge (q \vee r)$ is true and $(p \wedge q) \vee r$ is false. The second needs r false and $p \wedge q$ false; the first needs p true and $q \vee r$ true, so with r false, q true. But then $p \wedge q$ is true. No row makes the formula false, so it is a tautology (Definition 4.5.1).

Derivable without RAA. Completeness (Theorem 4.8.4) says it is derivable, but not how, and its proof uses LEM in Lemma 4.8.3, so it does not answer the second question. A search does: the goal is an implication, so a box on $p \wedge (q \vee r)$; split it by $\wedge$E and use $q \vee r$ by $\vee$E.

1	$p \wedge (q \vee r)$	asm, for $\to$I
2	p	$\wedge$E 1
3	$q \vee r$	$\wedge$E 1
4	q	asm, for $\vee$E
5	$p \wedge q$	$\wedge$I 2, 4
6	$(p \wedge q) \vee r$	$\vee$I 5
7	r	asm, for $\vee$E
8	$(p \wedge q) \vee r$	$\vee$I 7
9	$(p \wedge q) \vee r$	$\vee$E 3, 4–6, 7–8
10	$(p \wedge (q \vee r)) \to ((p \wedge q) \vee r)$	$\to$I 1–9

No RAA is used: the boxes close by $\vee$E and $\to$I. The tools: truth values (Definition 4.4.2) for the first question, the search of Chapter 3 for the second. The look-alike tool, completeness, answers only "derivable with all the rules".

4.9 Normal Forms and Adequate Connectives

The proof of completeness wrote each row of a truth table as a list of signed atoms $p_1^v, \ldots, p_n^v$ (Section 4.8). Join the signed atoms of one row by $\wedge$, and you get a formula that is true in that row and in no other. Join those formulas, one for each row where a given formula φ is true, by $\vee$, and you get a formula with the same truth table as φ. This section turns that observation into two standard shapes for formulas, the disjunctive and the conjunctive normal form. It then

reads validity off the conjunctive shape at a glance, and asks how few connectives are really needed.

On a first reading this section may wait: it is a second-pass section, and nothing in Chapters 5–8 uses it.

4.9.1 Conjunctions and disjunctions of many formulas

Normal forms are long conjunctions and disjunctions. Section 4.5 fixed how to write them, in the paragraph on long conjunctions after Proposition 4.5.1. For formulas $\chi_1, \dots, \chi_k$ with $k \geq 1$, $\chi_1 \wedge \chi_2 \wedge \cdots \wedge \chi_k$ is the conjunction **grouped to the right**, $\chi_1 \wedge (\chi_2 \wedge (\cdots \wedge \chi_k))$, and likewise for $\vee$; for $k = 1$ it is χ_1 itself. That paragraph proved, by induction on the number of terms, how such formulas are evaluated: for every valuation v,

$$\begin{aligned} v(\chi_1 \wedge \cdots \wedge \chi_k) = \mathrm{T} &\iff v(\chi_i) = \mathrm{T} \text{ for every } i = 1, \dots, k, \\ v(\chi_1 \vee \cdots \vee \chi_k) = \mathrm{T} &\iff v(\chi_i) = \mathrm{T} \text{ for some } i = 1, \dots, k. \end{aligned} \tag{4.9.1}$$

We use these two lines constantly below, and we cite them by this display.

So a long conjunction is true when **all** its parts are, and a long disjunction when **at least one** is, just as the words say.

4.9.2 Literals, clauses and the two normal forms

The building blocks are the smallest formulas that still carry a sign.

Definition 4.9.1. Literal

A **literal** is an atom or the negation of an atom.

So p, $\neg p$, q_7 and $\neg q_7$ are literals. The signed atom p^v of Section 4.8 is always a literal. The formulas $\neg\neg p$, $\neg(p \wedge q)$ and $\bot$ are not: in the first two, $\neg$ is applied to something other than an atom, and $\bot$ is neither an atom nor a negation. The **opposite** of a literal is $\neg p$ if the literal is p, and p if it is $\neg p$. By the clause for $\neg$ in Definition 4.4.2, a literal and its opposite take opposite values under every valuation.

From literals we build two kinds of blocks and two kinds of formulas.

- A **clause** joins one or more literals by $\vee$, such as $p \vee \neg q \vee r$, or $\neg q$ alone.
- A **conjunctive clause** joins one or more literals by $\wedge$, such as $p \wedge \neg q \wedge r$, or $\neg q$ alone. (We do not shorten it to "conjunct", which keeps its ordinary meaning, one side of a conjunction.)

A normal form is a formula whose connectives come in fixed layers: negations only on atoms, then one layer of one connective, then one layer of the other.

Definition 4.9.2. Disjunctive Normal Form

A formula **has disjunctive normal form** (DNF), or is **in DNF**, when it has the shape $C_1 \vee \cdots \vee C_k$, with $k \geq 1$, where every C_i joins **one or more literals** by $\wedge$.

Definition 4.9.3. Conjunctive Normal Form

A formula **has conjunctive normal form** (CNF), or is **in CNF**, when it has the shape $K_1 \wedge \cdots \wedge K_k$, with $k \geq 1$, where every K_i joins **one or more literals** by $\vee$.

In words: a DNF is an "or of ands" and a CNF an "and of ors", and in both the only negations sit directly on atoms. Read a DNF as a list of situations, any one of which makes it true; read a CNF as a list of requirements, all of which must be met.

Check some formulas against the definitions, clause by clause.

1. $(p \wedge \neg q) \vee (\neg p \wedge q)$ is in DNF: two conjunctive clauses, each a conjunction of literals. It is not in CNF: as a conjunction of clauses it would have to be a single clause (it is not a conjunction), and a clause cannot contain $\wedge$.
2. $(p \vee q) \wedge (\neg p \vee \neg q)$ is in CNF, with two clauses, and not in DNF.
3. $p \wedge \neg q$ is in both: in DNF with one conjunctive clause, and in CNF with two clauses of one literal each. Likewise $p \vee \neg q$ is in both.
4. A single literal, $\neg r$, is in both forms: one conjunctive clause of one literal, and one clause of one literal. This degenerate case matters, because a formula such as $\neg r$ must count as already normal.

A nonexample by a minimal change: $(p \wedge \neg q) \vee r$ is in DNF, but $\neg(p \wedge q) \vee r$ is not in either form. Only one symbol moved, and the literal condition fails: $\neg$ now stands on the conjunction $p \wedge q$, not on an atom. The formula $p \to q$ is in neither form either, since $\to$ is not allowed at all.

Why these layers? Because they are the shapes that the truth table hands us, as the next theorem shows, and because in these shapes some questions become easy: the subsection "Reading validity off a CNF" below decides validity of a CNF without a table.

Warning. A literal is an atom with **at most one** $\neg$, and nothing else. $\neg\neg p$ is equivalent to the literal p but is not a literal, so $\neg\neg p \vee q$ is not in CNF as written. Likewise $\bot$ is not a literal, so a "normal form" containing $\bot$ is not one; that is why the constructions below never use it.

Quick check. Which of the following are in CNF, and which are in DNF? (i) $(p \vee \neg q) \wedge r$; (ii) $\neg(p \vee q)$; (iii) $p \vee (\neg q \wedge r)$.

Solution. (i) CNF, with clauses $p \vee \neg q$ and r; not DNF, since the conjunctive clause would have to be the whole formula and $p \vee \neg q$ is not a literal. (ii) Neither: $\neg$ stands on a disjunction. (iii) DNF, with conjunctive clauses p and $\neg q \wedge r$; not CNF, since the one clause would contain $\wedge$.

4.9.3 Reading a DNF off the truth table

Take a valuation v and atoms $p_1, \dots, p_n$ with $n \geq 1$, and form the **minterm** of the row of v, the conjunctive clause of its signed atoms,

$$C_v := p_1^v \wedge \cdots \wedge p_n^v.$$

For every valuation w, each p_i^v is true under w exactly when $w(p_i) = v(p_i)$: if $v(p_i) = \mathrm{T}$, then $p_i^v = p_i$; if $v(p_i) = \mathrm{F}$, then $p_i^v = \neg p_i$, which is true under w exactly when $w(p_i) = \mathrm{F}$. By Eq. 4.9.1,

$$w(C_v) = \mathrm{T} \iff w(p_i) = v(p_i) \text{ for every } i = 1, \dots, n. \tag{4.9.2}$$

So C_v is true in the row of v and in no other row. It is a formula that "names" one row.

Example. Exactly one of two

Find a formula in DNF that is true exactly when exactly one of p, q is true, and check it.

Solution. The rows with exactly one of p, q true are $(p, q) = (\mathrm{T}, \mathrm{F})$ and (F, T). Their minterms are $p \wedge \neg q$ and $\neg p \wedge q$, so the answer is

$$(p \wedge \neg q) \vee (\neg p \wedge q).$$

Check against the table: in the row T, F the first minterm is true; in F, T the second is; in T, T each minterm has a false literal ($\neg q$, $\neg p$); in F, F likewise (p, q). So the formula is true in exactly the two rows wanted.

The same recipe works for any formula.

Theorem 4.9.1. Disjunctive Normal Form

Every formula φ is equivalent to a formula in DNF. If φ has at least one atom, the DNF can be chosen so that every atom in it is an atom of φ.

Idea. List the rows where φ is true and disjoin their minterms C_v. A valuation makes the disjunction true exactly when it agrees with one of those rows, which by Lemma 4.4.1 is exactly when it makes φ true. Two degenerate cases need a convention: a formula with no T-row, whose disjunction would be empty, and a formula with no atoms, which has no rows to speak of.

Proof. By the paraphrase of Definition 4.5.4, two formulas φ and ψ are equivalent exactly when $v(\varphi) = v(\psi)$ for every valuation v.

Case 1: φ has at least one atom. Let $p_1, \dots, p_n$ ($n \geq 1$) be the atoms of φ. A **row** is a choice of a value for each p_i; there are 2^n rows, and for each we fix a valuation that takes those values at $p_1, \dots, p_n$ (and, say, T at every other atom). By Lemma 4.4.1, $v(\varphi)$ depends only on the row of v. Let $v_1, \dots, v_k$ be the fixed valuations of the rows in which φ is T.

If $k \geq 1$, let $D \coloneqq C_{v_1} \vee \cdots \vee C_{v_k}$, which is in DNF by Definition 4.9.2, with atoms among $p_1, \ldots, p_n$. Let w be any valuation. By Eq. 4.9.1 and Eq. 4.9.2, $w(D) = \mathrm{T}$ exactly when w agrees with some v_j on $p_1, \ldots, p_n$. If it does, then $w(\varphi) = v_j(\varphi) = \mathrm{T}$ by Lemma 4.4.1. Conversely, if $w(\varphi) = \mathrm{T}$, then the row of w is one where φ is T, so w agrees on $p_1, \ldots, p_n$ with the valuation v_j fixed for that row. Hence $w(D) = w(\varphi)$ for every w, and $\varphi \equiv D$.

If $k = 0$, then φ is F in every row, hence under every valuation. The formula $p_1 \wedge \neg p_1$ is in DNF (one conjunctive clause of two literals) and is F under every valuation by Definition 4.4.2. So $\varphi \equiv p_1 \wedge \neg p_1$.

Case 2: φ has no atoms. Any two valuations agree on all the atoms of φ (there are none), so by Lemma 4.4.1 $v(\varphi)$ is the same for every v. Let p be an atom. If that value is T, then $\varphi \equiv p \vee \neg p$, which is true under every valuation and is in DNF (two conjunctive clauses of one literal); if it is F, then $\varphi \equiv p \wedge \neg p$. This proves the theorem.

■

The minterms of different rows are never true together, and in each row exactly one of the 2^n minterms is true: they divide the rows like the pieces of a partition, and a full disjunctive normal form collects the pieces in which the formula is true.

The minterms C_v are the signed atoms of Section 4.8 joined by $\wedge$. There the signed atoms of a row were premises from which φ^v was derived; here they are joined into a formula that is true in that row only. The proof is the proof of completeness with the derivations taken out.

↻ **Second reading** · *Why not the constant for contradiction*

For $p \wedge \neg p$ we could have written $\bot$, which is also F under every valuation. We did not, because $\bot$ is not a literal, and because the adequacy theorem at the end of this section reuses the construction in a language without $\bot$.

4.9.4 Conjunctive normal form

A DNF lists the rows where a formula is true. A CNF does the opposite: it lists the rows where the formula is **false**, one clause per row, each clause being false in its row and true in every other.

Theorem 4.9.2. Conjunctive Normal Form

Every formula φ is equivalent to a formula in CNF. If φ has at least one atom, the CNF can be chosen so that every atom in it is an atom of φ.

Idea. Apply Theorem 4.9.1 to $\neg\varphi$ and negate the result: $\varphi \equiv \neg\neg\varphi \equiv \neg(C_1 \vee \cdots \vee C_k)$. Now push the $\neg$ inwards with De Morgan's laws and double negation (Proposition 4.5.1): $\neg(C_1 \vee \cdots \vee C_k)$ becomes $\neg C_1 \wedge \cdots \wedge \neg C_k$, and each $\neg C_i$, the negation of a conjunction of literals, becomes the disjunction of the opposite literals. The proof checks the outcome directly with Eq. 4.9.1, which avoids De Morgan's laws for many terms.

Proof. By Theorem 4.9.1, $\neg\varphi \equiv C_1 \vee \cdots \vee C_k$ for some conjunctive clauses $C_1, \ldots, C_k$ (with atoms among those of φ if it has any, since $\neg\varphi$ has the same atoms). For each i, let K_i be the disjunction of the opposites of the literals of C_i, in the same order, and let $K \coloneqq K_1 \wedge \cdots \wedge K_k$. Then K is in CNF, with the same atoms as $C_1 \vee \cdots \vee C_k$.

Let w be a valuation. Each literal of K_i is true under w exactly when the corresponding literal of C_i is false. By Eq. 4.9.1, used at each step:

$$\begin{aligned} w(K) = \mathrm{T} &\iff \text{for every } i,\ w(K_i) = \mathrm{T} \\ &\iff \text{for every } i,\ \text{some literal of } C_i \text{ is false under } w \\ &\iff \text{for every } i,\ w(C_i) = \mathrm{F} \\ &\iff w(C_1 \vee \cdots \vee C_k) = \mathrm{F}. \end{aligned}$$

Since $C_1 \vee \cdots \vee C_k \equiv \neg\varphi$, the last line holds exactly when $w(\neg\varphi) = \mathrm{F}$, that is, when $w(\varphi) = \mathrm{T}$. So $w(K) = w(\varphi)$ for every w, and $\varphi \equiv K$.

■

Unwinding the construction: the conjunctive clauses C_i of the DNF of $\neg\varphi$ are the rows where φ is **false**, and K_i is the clause that is false exactly in the row of C_i. So the CNF can be read straight off the table: **one clause per F-row**, built from the opposites of that row's signed atoms. In the row $p = \mathrm{T}$, $q = \mathrm{F}$, for instance, the clause is $\neg p \vee q$. A formula with no F-row gets, by the proof, the clause of a contradiction's DNF negated, $\neg p \vee p$ (from $C_1 = p \wedge \neg p$), which is true in every row.

Quick check. On the atoms p, q, r, write the CNF clause for the F-row $p = \mathrm{F}$, $q = \mathrm{T}$, $r = \mathrm{F}$.

Solution. The row's signed atoms are $\neg p$, q, $\neg r$; their opposites give $p \vee \neg q \vee r$. It is false exactly when p is F, q is T and r is F, that is, in that row only.

Alternatively, a CNF can be reached by **rewriting**. First replace each $\bot$, which is not a literal, by $p \wedge \neg p$ for an atom p; both are false in every row. Replace $\alpha \leftrightarrow \beta$ by $(\alpha \to \beta) \wedge (\beta \to \alpha)$ and $\alpha \to \beta$ by $\neg\alpha \vee \beta$; push each $\neg$ inwards by De Morgan's laws, and remove $\neg\neg$; finally distribute $\vee$ over $\wedge$, replacing $\alpha \vee (\beta \wedge \gamma)$ by $(\alpha \vee \beta) \wedge (\alpha \vee \gamma)$. Each step replaces a part of the formula by an equivalent one: the equivalences are $\bot \equiv p \wedge \neg p$ and those of Proposition 4.5.1, which hold for arbitrary formulas, and Lemma 4.5.2 says the whole formula stays equivalent; since $\equiv$ is transitive (Section 4.5), the steps chain. Stopping before the distribution step, and distributing $\wedge$ over $\vee$ instead, gives a DNF. The rewriting route is often much shorter than the table, as the next example shows, but it can also explode, as the warning after it shows.

Example. Two routes to a normal form

Put $\varphi = (p \to q) \to (\neg q \to r)$ into DNF and into CNF, first from its truth table and then by rewriting.

Solution. The table route. Work backwards from F (Section 4.4): φ is false exactly when $p \to q$ is T and $\neg q \to r$ is F. The latter needs $\neg q$ true and r false, so $q = r = \mathrm{F}$; then $p \to q$ true needs $p = \mathrm{F}$. So φ has exactly one F-row, $p = q = r = \mathrm{F}$, as the table confirms:

p	q	r	$p \to q$	$\neg q \to r$	$(p \to q) \to (\neg q \to r)$
T	T	T	T	T	T
T	T	F	T	T	T
T	F	T	F	T	T
T	F	F	F	F	T
F	T	T	T	T	T
F	T	F	T	T	T
F	F	T	T	T	T
F	F	F	T	F	F

The DNF of Theorem 4.9.1 has one minterm for each of the seven T-rows:

$$(p \wedge q \wedge r) \vee (p \wedge q \wedge \neg r) \vee (p \wedge \neg q \wedge r) \vee (p \wedge \neg q \wedge \neg r) \\ \vee (\neg p \wedge q \wedge r) \vee (\neg p \wedge q \wedge \neg r) \vee (\neg p \wedge \neg q \wedge r).$$

The CNF of Theorem 4.9.2 has one clause for the single F-row, built from the opposites of $\neg p, \neg q, \neg r$: it is the single clause $p \vee q \vee r$.

The rewriting route. Eliminate the three implications, then push the negation in:

$$\begin{aligned} \varphi &\equiv \neg(\neg p \vee q) \vee (\neg\neg q \vee r) \\ &\equiv (\neg\neg p \wedge \neg q) \vee (q \vee r) \\ &\equiv (p \wedge \neg q) \vee q \vee r, \end{aligned}$$

using (c) of Proposition 4.5.1 three times, then De Morgan's law (b) for $\neg(\alpha \vee \beta)$, then double negation (d) twice. This is already a DNF, with three conjunctive clauses. For a CNF, distribute, by (g) together with commutativity (f) and associativity (h): $(p \wedge \neg q) \vee (q \vee r) \equiv (p \vee q \vee r) \wedge (\neg q \vee q \vee r)$. The second clause contains q and $\neg q$, so it is true in every row, and a conjunction with a part that is true in every row is equivalent to its other part. So $\varphi \equiv p \vee q \vee r$: one clause, as the table route found.

The formula $p \vee q \vee r$ is also a DNF, with three conjunctive clauses of one literal each. So the table's seven minterms and the rewriting's three conjunctive clauses can both be replaced by three one-literal conjunctive clauses. Both routes are correct; they differ in length, and neither is guaranteed to give the shortest normal form.

Warning. A formula has **many** equivalent CNFs and DNFs; "normal form" does not mean "unique form". $p \vee q \vee r$ and $(p \vee q \vee r) \wedge (\neg q \vee q \vee r)$ are both CNFs of the formula above. And the two routes have opposite weaknesses. The table route needs all 2^n rows for n atoms. The rewriting route can explode at the distribution step: distributing $(p_1 \wedge q_1) \vee (p_2 \wedge q_2) \vee (p_3 \wedge q_3)$ gives 8 clauses, one for each way of picking p_i or q_i from every conjunctive clause, and m such conjunctive clauses give 2^m clauses.

4.9.5 Reading validity off a CNF

A CNF is a list of requirements, all of which must hold. That makes one question easy.

Proposition 4.9.3. Validity of a CNF

(a) A formula in CNF is a tautology exactly when **every** clause contains some atom together with its negation.

(b) A formula in DNF is satisfiable exactly when **some** conjunctive clause contains no atom together with its negation.

Try it first (ten minutes): take a clause in which no atom occurs both with and without $\neg$, such as $p \vee \neg q \vee r$. Can you find a row that makes it false? What does that row do to a CNF containing this clause?

Idea. The direction "every clause has a pair $\Rightarrow$ tautology" is the easy one: $p \vee \neg p$ is true in every row, so a clause containing both is too, and a conjunction of clauses true in every row is true in every row. The attempt above is the other direction. To make a CNF false, it is enough to make **one** clause false, and a clause is false only when **all** its literals are. So set every atom of the clause against its literal: F for a plain atom, T for a negated one. The one thing that could stop us is an atom that appears both plain and negated, since it would have to be both F and T. That is exactly the pair the proposition talks about.

Proof. *(a)* Let $K = K_1 \wedge \cdots \wedge K_k$ be in CNF.

($\Leftarrow$) Suppose every clause K_i contains some atom p and its negation $\neg p$. Let v be a valuation. One of p, $\neg p$ is true under v, by Definition 4.4.2, so $v(K_i) = \mathrm{T}$ by Eq. 4.9.1. This holds for every i, so $v(K) = \mathrm{T}$ by Eq. 4.9.1. As v was arbitrary, K is a tautology (Definition 4.5.1).

($\Rightarrow$) We prove the contrapositive. Suppose some clause K_i contains no atom together with its negation. Define a valuation v by: $v(p) := \mathrm{F}$ if the literal p occurs in K_i; $v(p) := \mathrm{T}$ if the literal $\neg p$ occurs in K_i; $v(p) := \mathrm{T}$ for every other atom. This is well defined, because no atom occurs in K_i both as p and as $\neg p$. Every literal of K_i is false under v, so $v(K_i) = \mathrm{F}$ by Eq. 4.9.1, and hence $v(K) = \mathrm{F}$, again by Eq. 4.9.1. So K is not a tautology.

(b) Let $D = C_1 \vee \cdots \vee C_k$ be in DNF.

($\Leftarrow$) Suppose C_i contains no atom together with its negation. Define v by $v(p) := \mathrm{T}$ if p occurs in C_i, $v(p) := \mathrm{F}$ if $\neg p$ occurs in C_i, and $v(p) := \mathrm{T}$ otherwise; this is well defined as in (a). Every literal of C_i is true under v, so $v(C_i) = \mathrm{T}$ and $v(D) = \mathrm{T}$ by Eq. 4.9.1. So D is satisfiable (Definition 4.5.2).

($\Rightarrow$) If every conjunctive clause contains some p and $\neg p$, then under every valuation one of these two literals is false, so every conjunctive clause is false and so is D, by Eq. 4.9.1. So D is not satisfiable. This proves the proposition.

■

Second reading · *Two faces of one fact*

The two parts are two faces of one fact. Section 4.5 showed that a formula is a tautology exactly when its negation is unsatisfiable (after Definition 4.5.2). Let D be in DNF, and let K be the CNF whose clauses are the conjunctive clauses of D with every literal replaced by its opposite. The proof of Theorem 4.9.2, applied to $\varphi = \neg D$ with D itself as the DNF of $\neg\varphi \equiv D$, shows $K \equiv \neg D$. A clause of K has an atom together with its negation exactly when the corresponding conjunctive clause of D has. So D is unsatisfiable exactly when $\neg D$, and so K, is a tautology, which by (a) happens exactly when every conjunctive clause of D has such a pair. That is (b) again.

The proposition trades one kind of work for another. Checking a CNF for validity is now a glance, but getting the CNF in the first place may take the 2^n rows of the table route or the explosion of the rewriting route. Nothing is free: the work has moved, not vanished.

Example. Validity at a glance

Decide whether each formula is a tautology, without a table.

(a) $(\neg p \vee q \vee p) \wedge (r \vee \neg q \vee s \vee q) \wedge (\neg s \vee \neg r \vee s) \wedge (p \vee \neg r \vee r \vee q)$.

(b) The same formula with the third clause changed to $\neg s \vee \neg r \vee q$.

Solution. (a) Each clause has a pair: $\neg p, p$ in the first; $\neg q, q$ in the second; $\neg s, s$ in the third; $\neg r, r$ in the fourth. By Proposition 4.9.3 (a), it is a tautology. A table would have needed 16 rows.

(b) The new third clause $\neg s \vee \neg r \vee q$ has the atoms s, r, q, each once, so it has no pair, and the formula is not a tautology. The proof of Proposition 4.9.3 names a refuting row: $s = \mathrm{T}$, $r = \mathrm{T}$, $q = \mathrm{F}$, and $p = \mathrm{T}$, like every atom not in the clause; in fact either value of p works. In that row the third clause is F, so the conjunction is F.

4.9.6 Calculating with equivalences

Section 4.5 replaced a part of a formula by an equivalent part (Lemma 4.5.2) and chained such steps, since $\equiv$ is transitive. Normal forms can be reached in this way, as the rewriting route did, and so can shorter formulas. Each step cites a law, as each line of a derivation cites a rule. A few more laws make the calculations run; in them, $\neg\bot$ serves as a formula that is true in every row.

Proposition 4.9.4. Laws of the Constants, Idempotence and Absorption

For all formulas φ and ψ:

(a) $\varphi \wedge \neg\bot \equiv \varphi$ and $\varphi \vee \bot \equiv \varphi$;

(b) $\varphi \wedge \bot \equiv \bot$ and $\varphi \vee \neg\bot \equiv \neg\bot$;

(c) $\varphi \wedge \neg\varphi \equiv \bot$ and $\varphi \vee \neg\varphi \equiv \neg\bot$;

(d) $\varphi \wedge \varphi \equiv \varphi$ and $\varphi \vee \varphi \equiv \varphi$;
(e) $\varphi \wedge (\varphi \vee \psi) \equiv \varphi$ and $\varphi \vee (\varphi \wedge \psi) \equiv \varphi$.

↻ **Second reading** · *The tables for the further laws*

Proof. As for Proposition 4.5.1: fix a valuation v, and let $a = v(\varphi)$ and $b = v(\psi)$. By (V1) and (V2) of Definition 4.4.2, $v(\bot) = \mathrm{F}$ and $v(\neg\bot) = \mathrm{T}$, so the value of each side under v is computed from a (and b) by the clauses of its connectives. The first table covers (a)–(d), with one row for each value of a; the second covers (e), with one row for each pair (a, b).

a	$\varphi \wedge \neg\bot$	$\varphi \vee \bot$	$\varphi \wedge \bot$	$\varphi \vee \neg\bot$	$\varphi \wedge \neg\varphi$	$\varphi \vee \neg\varphi$	$\varphi \wedge \varphi$	$\varphi \vee \varphi$
T	T	T	F	T	F	T	T	T
F	F	F	F	T	F	T	F	F

a	b	$\varphi \wedge (\varphi \vee \psi)$	$\varphi \vee (\varphi \wedge$
T	T	T	T
T	F	T	T
F	T	F	F
F	F	F	F

In (a), (d) and (e) both columns agree with a; in (b) and (c) the columns are constant, F where the right side is $\bot$ and T where it is $\neg\bot$. So both sides of each law have the same value under v, and as v was arbitrary they are equivalent, by the paraphrase of Definition 4.5.4.

■

Example. Simplifying by a chain

Simplify $(p \to q) \wedge (p \to \neg q)$, citing a law at each step.

Solution.

$$\begin{aligned}(p \to q) \wedge (p \to \neg q) &\equiv (\neg p \vee q) \wedge (\neg p \vee \neg q)\\ &\equiv \neg p \vee (q \wedge \neg q)\\ &\equiv \neg p \vee \bot\\ &\equiv \neg p.\end{aligned}$$

Step 1 is (c) of Proposition 4.5.1, used twice, each time applied inside the formula by Lemma 4.5.2. Step 2 is (g), second form, read from right to left with $\varphi = \neg p$, $\psi = q$, $\chi = \neg q$. Step 3 is (c) of Proposition 4.9.4, applied inside. Step 4 is (a).

Read the result: if p gives both q and not q, then not p. This is the table's face of ¬I.

Quick check. Simplify $(p \vee q) \wedge (p \vee \neg q)$, citing a law at each step.

Solution. $(p \vee q) \wedge (p \vee \neg q) \equiv p \vee (q \wedge \neg q) \equiv p \vee \bot \equiv p$: (g) of Proposition 4.5.1, second form, read from right to left; then (c) of Proposition 4.9.4, first form, applied inside by

Lemma 4.5.2; then its (a), second form.

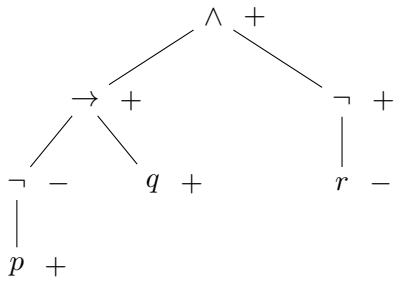
4.9.7 Stronger and weaker parts

Replacing a part by an equivalent part keeps the value of the whole (Lemma 4.5.2). Replacing it by a stronger or a weaker part moves the value in one direction, and which direction depends on where the part sits.

Definition 4.9.4. Positive and Negative Occurrences

Each node of the parse tree of a formula χ (an **occurrence** of a subformula, Section 4.5) is **positive**, **negative**, or **unsigned**, by the path from the root: the root is positive; the part of a node $\neg\theta$, and the left part of a node $\theta_1 \to \theta_2$, have the sign opposite to that node; the parts of a node $\theta_1 \wedge \theta_2$ or $\theta_1 \vee \theta_2$, and the right part of $\theta_1 \to \theta_2$, have the same sign as that node; the parts of a node $\theta_1 \leftrightarrow \theta_2$, and every node below them, are unsigned.

In words: walk down from the root; the sign flips at each negation and at each hypothesis of an implication, and stays at every other step. A biconditional uses each part in both directions, as hypothesis and as conclusion, so below it nothing is signed. The sign is well defined because a formula has only one parse tree (Section 4.3).



The parse tree of $(\neg p \to q) \wedge \neg r$, each node with its sign. Signs flip at each $\neg$ and on the left of each $\to$: here p is positive (two flips), q positive, r negative.

Theorem 4.9.5. Monotonicity

Let $\alpha \models \beta$, and let χ' be obtained from χ by replacing one occurrence of α by β. If the occurrence is positive, then $\chi \models \chi'$; if it is negative, then $\chi' \models \chi$.

In words: a stronger part in a positive place makes the whole stronger; in a negative place, weaker. Two readings are used daily. **A stronger hypothesis gives a weaker theorem**: the hypothesis of $\alpha \to \theta$ is negative, so from $\alpha \models \beta$ we get $(\beta \to \theta) \models (\alpha \to \theta)$. **It suffices to prove something stronger**: the conclusion is positive, so from $\alpha \models \beta$ we get $(\theta \to \alpha) \models (\theta \to \beta)$. For example, from $p \wedge q \models p$ the theorem gives $(r \to (p \wedge q)) \models (r \to p)$ and $(p \to r) \models ((p \wedge q) \to r)$.

Idea. Induct on χ, as for Lemma 4.5.2. There the replaced part kept its value, and so did the whole. Here truth passes one way only: from the old part to the new one, and up through $\wedge$, $\vee$ and the right of $\to$ unchanged; at a $\neg$ and at the left of a $\to$ the direction turns round, exactly where the sign flips.

Second reading · *The proof of monotonicity*

Proof. We apply Theorem 4.3.1 to the property $M(\chi)$: *for every positive or negative occurrence of α in χ, with χ' the result of replacing it by β, and every valuation v: if the occurrence is positive, $v(\chi) = \mathrm{T}$ implies $v(\chi') = \mathrm{T}$; if it is negative, $v(\chi') = \mathrm{T}$ implies $v(\chi) = \mathrm{T}$.* By Definition 4.5.3, these say $\chi \models \chi'$ and $\chi' \models \chi$.

Signs inside a part. Let θ be an immediate part of χ. The parse tree of θ is a subtree of that of χ (Definition 4.1.2), and the path from the root of χ to a node of θ is the path from the root of θ with one step on top. So the node's sign in χ is its sign in θ if θ is positive in χ, the opposite sign if θ is negative, and none if θ is unsigned.

If the occurrence is the root, it is positive, $\chi = \alpha$ and $\chi' = \beta$, and $v(\alpha) = \mathrm{T}$ implies $v(\beta) = \mathrm{T}$ because $\alpha \models \beta$. An atom or $\bot$ has no other node. Otherwise the occurrence lies in one immediate part, and χ' is χ with that part θ replaced by the result θ' of the replacement inside it.

$\chi = \neg\theta$. The part is negative, so the occurrence has opposite signs in χ and in θ. If it is positive in χ, it is negative in θ, and $M(\theta)$ gives: $v(\theta') = \mathrm{T}$ implies $v(\theta) = \mathrm{T}$. So if $v(\neg\theta) = \mathrm{T}$, then $v(\theta) = \mathrm{F}$, hence $v(\theta') = \mathrm{F}$, and $v(\neg\theta') = \mathrm{T}$ by (V2). If it is negative in χ, the same argument runs with θ and θ' exchanged.

$\chi = \theta_1 \wedge \theta_2$ *or* $\theta_1 \vee \theta_2$. Both parts are positive, so the signs in χ and in the part agree. Clauses (V3) and (V4) are monotone in each part: if a part is replaced by one that is true whenever it was, a true conjunction or disjunction stays true. With M of the part, this gives the claim in both cases.

$\chi = \theta_1 \to \theta_2$. In θ_2 the signs agree, and (V5) is monotone in the conclusion: if $v(\theta_2) = \mathrm{T}$ implies $v(\theta_2') = \mathrm{T}$, then $v(\theta_1 \to \theta_2) = \mathrm{T}$ implies $v(\theta_1 \to \theta_2') = \mathrm{T}$. In θ_1 the signs are opposite, and (V5) is antitone in the hypothesis: if $v(\theta_1') = \mathrm{T}$ implies $v(\theta_1) = \mathrm{T}$, then $v(\theta_1 \to \theta_2) = \mathrm{T}$ implies $v(\theta_1' \to \theta_2) = \mathrm{T}$, since a false θ_1 leaves θ_1' false. An occurrence positive in χ is negative in θ_1, so $M(\theta_1)$ supplies exactly this premise; a negative one is positive in θ_1, and the argument runs with θ_1 and θ_1' exchanged.

$\chi = \theta_1 \leftrightarrow \theta_2$. Every node below the root is unsigned, so there is no occurrence to treat.

By Theorem 4.3.1, $M(\chi)$ holds for every formula χ. This proves the theorem. ∎

Quick check. Give the sign of each atom occurrence in $\neg(p \to q) \vee r$.

Solution. p is positive: flipped by the $\neg$, and again as the hypothesis of $\to$. q is negative: flipped by the $\neg$ only. r is positive.

Warning. Below a $\leftrightarrow$ nothing may be strengthened. $p \wedge q \models p$, but $(p \wedge q) \leftrightarrow r$ and $p \leftrightarrow r$ entail each other in neither direction: the row $p = \mathrm{T}$, $q = r = \mathrm{F}$ makes the first true and the second false, and the row $p = r = \mathrm{T}$, $q = \mathrm{F}$ does the reverse.

4.9.8 Adequate sets of connectives

Our formulas use five connectives and $\bot$. Chapter 2 already showed that some of them can be traded for others: an implication for a disjunction (Proposition 2.5.6), a biconditional for two implications (Proposition 2.4.1). How far can this go? Before we can answer, we need to say what a set of connectives must be able to "say".

Every formula whose atoms are among $p_1, \ldots, p_n$ has a truth table: a value for each of the 2^n rows. Turn this around, and ask whether **every** table is the table of some formula.

Section 4.4 called such a table a **truth function** (Definition 4.4.4): a value for each list of inputs, with no formula attached; a formula **expresses** it when it has that table.

There are 2^n lists of inputs and two choices of output for each, so there are 2^{2^n} truth functions with n inputs: 4 with one input and 16 with two. Examples: the **negation** function with one input, $\mathrm{T} \mapsto \mathrm{F}$, $\mathrm{F} \mapsto \mathrm{T}$, expressed by $\neg p_1$; the two-input function that is T only at (T, T), expressed by $p_1 \wedge p_2$; the three-input **majority** function, T when at least two inputs are T, expressed by $(p_1 \wedge p_2) \vee (p_1 \wedge p_3) \vee (p_2 \wedge p_3)$. A degenerate case: the one-input function that is always T is expressed by $p_1 \vee \neg p_1$, and also by $p_2 \to p_2$, which does not mention p_1 at all. The definition allows that.

We say a formula is **in** a set S of connectives if every connective occurring in it belongs to S, and $\bot$ occurs in it only if $\bot$ belongs to S (for this purpose $\bot$ counts as a connective with no inputs).

Definition 4.9.5. Adequate Set of Connectives

A set S of connectives is **adequate** if **every** truth function, with **any** number $n \geq 1$ of inputs, is expressed by some formula in S.

A set containing an adequate set is adequate, since a formula in the smaller set is also in the larger one. The theorem below shows that $\{\neg, \wedge, \vee\}$ is adequate, so the full set $\{\neg, \wedge, \vee, \to, \leftrightarrow, \bot\}$ is too. The question is which smaller sets are.

Rewriting "or" with "not" and "and". The DNF of a truth table uses $\neg$, $\wedge$ and $\vee$. To do without $\vee$, rewrite each disjunction by De Morgan's law: $\alpha \vee \beta$ has the same table as $\neg(\neg\alpha \wedge \neg\beta)$. Done everywhere, from the leaves up, this is a translation t_1 of the formulas built from atoms with $\neg$, $\wedge$, $\vee$ (the proof below gives the clauses for $\bot$, $\to$, $\leftrightarrow$): $t_1(p) = p$, $t_1(\neg\alpha) = \neg t_1(\alpha)$, $t_1(\alpha \wedge \beta) = t_1(\alpha) \wedge t_1(\beta)$, $t_1(\alpha \vee \beta) = \neg(\neg t_1(\alpha) \wedge \neg t_1(\beta))$. For instance

$$t_1\big((p \vee q) \wedge \neg r\big) = \neg(\neg p \wedge \neg q) \wedge \neg r.$$

In the row $p = \mathrm{F}$, $q = \mathrm{T}$, $r = \mathrm{F}$ both sides are T: on the left $p \vee q$ and $\neg r$ are T; on the right $\neg p \wedge \neg q$ is F, so its negation is T, and $\neg r$ is T. Two more translations do the same for the other sets of the theorem below:

	t_1, into $\{\neg, \wedge\}$	t_2, into $\{\neg, \vee\}$	t_3, into $\{\to, \bot\}$
$\neg\alpha$	$\neg t_1(\alpha)$	$\neg t_2(\alpha)$	$t_3(\alpha) \to \bot$
$\alpha \wedge \beta$	$t_1(\alpha) \wedge t_1(\beta)$	$\neg(\neg t_2(\alpha) \vee \neg t_2(\beta))$	$(t_3(\alpha) \to (t_3(\beta) \to \bot)) \to \bot$
$\alpha \vee \beta$	$\neg(\neg t_1(\alpha) \wedge \neg t_1(\beta))$	$t_2(\alpha) \vee t_2(\beta)$	$(t_3(\alpha) \to \bot) \to t_3(\beta)$

Quick check. Use the translation t_3 to express $p \wedge q$ with $\to$ and $\bot$ only, and check your formula in the row $p = \mathrm{T}$, $q = \mathrm{F}$.

Solution. $t_3(p \wedge q) = (p \to (q \to \bot)) \to \bot$. In the row $p = \mathrm{T}$, $q = \mathrm{F}$: $q \to \bot$ is $\mathrm{F} \to \mathrm{F}$, so T; then $p \to (q \to \bot)$ is T; then the whole is $\mathrm{T} \to \bot$, so F, which agrees with $p \wedge q$.

Theorem 4.9.6. Adequate Sets

(a) The sets $\{\neg, \wedge, \vee\}$, $\{\neg, \wedge\}$, $\{\neg, \vee\}$ and $\{\to, \bot\}$ are adequate.

(b) The set $\{\wedge, \vee, \to\}$ is not adequate: no formula in it expresses the negation function.

Idea. For (a), the construction of Theorem 4.9.1 works for any table, not only for the table of a formula, and its output uses only $\neg$, $\wedge$, $\vee$. To shrink the set further, rewrite each missing connective in terms of the others, for instance $\alpha \vee \beta$ as $\neg(\neg\alpha \wedge \neg\beta)$, recursively through the formula. For (b), look at the row where every atom is true: $\wedge$, $\vee$ and $\to$ all give T on T inputs, so every formula in them is T there, while negation must give F.

Proof. *(a) The set $\{\neg, \wedge, \vee\}$.* Let f be a truth function with $n \geq 1$ inputs. For each of the 2^n lists of inputs, fix a valuation taking those values at $p_1, \ldots, p_n$, and let $v_1, \ldots, v_k$ be those fixed for the lists where f is T. If $k \geq 1$, let $D := C_{v_1} \vee \cdots \vee C_{v_k}$, with the minterms C_v formed from $p_1, \ldots, p_n$; if $k = 0$, let $D := p_1 \wedge \neg p_1$. By Eq. 4.9.1 and Eq. 4.9.2, a valuation w makes D true exactly when $(w(p_1), \ldots, w(p_n))$ is a list where f is T, that is, when $f(w(p_1), \ldots, w(p_n)) = \mathrm{T}$. So D expresses f, and D is in $\{\neg, \wedge, \vee\}$.

The other three sets. By Theorem 4.3.4, the following clauses define three functions t_1, t_2, t_3 on all formulas. Each has $t_i(p) := p$ for every atom p, and $t_i(\bot) := \bot$, $t_i(\alpha \to \beta) := t_i(\alpha) \to t_i(\beta)$ and $t_i(\alpha \leftrightarrow \beta) := t_i(\alpha) \leftrightarrow t_i(\beta)$ (these last clauses are never used below). The other clauses rewrite the connectives $\neg$, $\wedge$, $\vee$:

- t_1, into $\{\neg, \wedge\}$: $t_1(\neg\alpha) := \neg t_1(\alpha)$, $t_1(\alpha \wedge \beta) := t_1(\alpha) \wedge t_1(\beta)$, and $t_1(\alpha \vee \beta) := \neg(\neg t_1(\alpha) \wedge \neg t_1(\beta))$;
- t_2, into $\{\neg, \vee\}$: $t_2(\neg\alpha) := \neg t_2(\alpha)$, $t_2(\alpha \wedge \beta) := \neg(\neg t_2(\alpha) \vee \neg t_2(\beta))$, and $t_2(\alpha \vee \beta) := t_2(\alpha) \vee t_2(\beta)$;
- t_3, into $\{\to, \bot\}$: $t_3(\neg\alpha) := t_3(\alpha) \to \bot$, $t_3(\alpha \wedge \beta) := (t_3(\alpha) \to (t_3(\beta) \to \bot)) \to \bot$, and $t_3(\alpha \vee \beta) := (t_3(\alpha) \to \bot) \to t_3(\beta)$.

Claim. Let $S_1 = \{\neg, \wedge\}$, $S_2 = \{\neg, \vee\}$, $S_3 = \{\to, \bot\}$. For every formula φ in $\{\neg, \wedge, \vee\}$ and each i, the formula $t_i(\varphi)$ is in S_i, and $v(t_i(\varphi)) = v(\varphi)$ for every valuation v.

Proof of Claim. By structural induction (Theorem 4.3.1) on φ, with the property "if φ is in $\{\neg, \wedge, \vee\}$, then for each i, $t_i(\varphi)$ is in S_i and has the same value as φ under every valuation". The immediate parts of a formula in $\{\neg, \wedge, \vee\}$ are in $\{\neg, \wedge, \vee\}$, so the induction hypothesis applies to them. For $\bot$, $\alpha \to \beta$ and $\alpha \leftrightarrow \beta$ the property holds vacuously, since these formulas are not in $\{\neg, \wedge, \vee\}$. For an atom, $t_i(p) = p$. In the other cases, write $a = v(t_i(\alpha)) = v(\alpha)$ and $b = v(t_i(\beta)) = v(\beta)$, by the induction hypothesis. Each clause for t_i above uses only the connectives of S_i, applied to formulas in S_i, so $t_i(\varphi)$ is in S_i. For the values, use Definition 4.4.2, and that $\chi \to \bot$ is true exactly when χ is false ($\bot$ is always F):

- $\neg t_i(\alpha)$ and $t_3(\alpha) \to \bot$ are true exactly when $a = \mathrm{F}$, as is $\neg\alpha$.
- $\neg(\neg t_2(\alpha) \vee \neg t_2(\beta))$ is true exactly when neither $\neg t_2(\alpha)$ nor $\neg t_2(\beta)$ is, that is, when $a = b = \mathrm{T}$. And $(t_3(\alpha) \to (t_3(\beta) \to \bot)) \to \bot$ is true exactly when $t_3(\alpha) \to (t_3(\beta) \to \bot)$ is false, that is, when $a = \mathrm{T}$ and $t_3(\beta) \to \bot$ is false, that is, when $a = b = \mathrm{T}$. Both agree with $\alpha \wedge \beta$.
- $\neg(\neg t_1(\alpha) \wedge \neg t_1(\beta))$ is true exactly when a and b are not both F. And $(t_3(\alpha) \to \bot) \to t_3(\beta)$ is true exactly when $t_3(\alpha) \to \bot$ is false or $b = \mathrm{T}$, that is, when $a = \mathrm{T}$ or $b = \mathrm{T}$. Both agree with $\alpha \vee \beta$.

The remaining clauses, for $t_1(\alpha \wedge \beta)$ and $t_2(\alpha \vee \beta)$, keep the connective and have the right value by the induction hypothesis.

■

Now let f be any truth function, and D the formula in $\{\neg, \wedge, \vee\}$ that expresses it. By the Claim, $t_i(D)$ is in S_i and $v(t_i(D)) = v(D) = f(v(p_1), \ldots, v(p_n))$ for every valuation v. So each S_i is adequate.

(b) Let v_{T} be the valuation with $v_{\mathrm{T}}(p) = \mathrm{T}$ for every atom p. We show by structural induction that every formula in $\{\wedge, \vee, \to\}$ is true under v_{T}; the property is "if φ is in $\{\wedge, \vee, \to\}$, then $v_{\mathrm{T}}(\varphi) = \mathrm{T}$". For an atom, $v_{\mathrm{T}}(p) = \mathrm{T}$. The formulas $\bot$, $\neg\alpha$ and $\alpha \leftrightarrow \beta$ are not in the set, so for them the property holds vacuously. If φ is $\alpha \wedge \beta$, $\alpha \vee \beta$ or $\alpha \to \beta$ and is in the set, then so are α and β, and by the induction hypothesis $v_{\mathrm{T}}(\alpha) = v_{\mathrm{T}}(\beta) = \mathrm{T}$. By Definition 4.4.2, $v_{\mathrm{T}}(\alpha \wedge \beta) = \mathrm{T}$ and $v_{\mathrm{T}}(\alpha \vee \beta) = \mathrm{T}$; and $v_{\mathrm{T}}(\alpha \to \beta) = \mathrm{T}$ because its conclusion β is true. By Theorem 4.3.1, every formula in the set is true under v_{T}.

Now suppose some φ in $\{\wedge, \vee, \to\}$ expressed the negation function f. Then $v_{\mathrm{T}}(\varphi) = f(v_{\mathrm{T}}(p_1)) = f(\mathrm{T}) = \mathrm{F}$, contradicting $v_{\mathrm{T}}(\varphi) = \mathrm{T}$. So no such φ exists, and the set is not adequate. This proves the theorem.

■

In the implication case of (b), the reason is that the **conclusion** β is true. That is where a careless argument goes wrong, and Exercise (C2: Find the error, used later) asked you to find a "proof" that goes wrong there.

The five connectives were chosen because each matches a word mathematicians use, not because they are needed: two suffice, and Exercise (C1: One connective is enough) shows that one well-chosen connective suffices.

Warning. The $\bot$ in $\{\to, \bot\}$ is essential. Without it, $\{\to\}$ is not adequate: every formula in $\{\to\}$ is also in $\{\wedge, \vee, \to\}$, so by (b) of Theorem 4.9.6 none expresses negation. Implication alone cannot say "false"; implication together with a formula that is always false can.

4.9.9 Exercises

4.9.9.1 A. Check your understanding

Exercise. A1

(a) Which of $\neg p$, $\neg\neg q$, $\bot$, r and $\neg(p \vee q)$ are literals?

(b) Classify each as a clause, a conjunctive clause, both, or neither: $p \vee \neg q$; $\neg r$; $p \wedge (q \vee r)$.

(c) Is $(p \vee q) \wedge (\neg p \vee p \vee r)$ a tautology? Is $(q \vee \neg q) \wedge (\neg r \vee p \vee r)$? Give a refuting row where the answer is no.

(d) Give a formula in DNF that is not satisfiable.

(e) True or false: every formula in CNF with only one clause is also in DNF.

Solution. *(a)* $\neg p$ and r. $\neg\neg q$ has two negations, $\bot$ is not an atom, and $\neg(p \vee q)$ negates a disjunction.

(b) $p \vee \neg q$ is a clause. $\neg r$ is both (one literal). $p \wedge (q \vee r)$ is neither: its part $q \vee r$ is not a literal.

(c) The first is not a tautology: its first clause $p \vee q$ has no pair, and the row $p = q = \mathrm{F}$ (any r) makes it false, hence the whole conjunction. The second is: each clause has a pair ($q, \neg q$ and $r, \neg r$), by Proposition 4.9.3 (a).

(d) $p \wedge \neg p$: its one conjunctive clause contains a pair, so by Proposition 4.9.3 (b) it is not satisfiable.

(e) True. A single clause $\ell_1 \vee \cdots \vee \ell_m$ is a disjunction of m conjunctive clauses of one literal each.

Exercise. A2: Drill

The parts run from one step of a calculation, through signs and what monotonicity gives, to normal forms and an adequate set.

(a) Simplify $p \wedge (p \vee q)$.

(b) Simplify $\neg(p \to q)$.

(c) Simplify $(p \vee q) \wedge \neg p$.

(d) Simplify $(p \wedge q) \vee (p \wedge \neg q)$.

(e) What is the sign of p in $\neg(p \to q)$?

(f) What is the sign of q in $(p \to \neg q) \wedge r$?

(g) What is the sign of p in $p \leftrightarrow q$?

(h) From $p \wedge q \models p$, what does Theorem 4.9.5 give for $r \to (p \wedge q)$?

(i) Give a DNF of $p \leftrightarrow q$.

(j) Give a CNF of $p \leftrightarrow q$.

(k) How many minterms has the DNF of Theorem 4.9.1 for a tautology whose atoms are p, q, r?

(l) Is $\{\neg, \to\}$ adequate?

Solution.

(a) p, by Proposition 4.9.4 (e).

(b) $p \wedge \neg q$: (c), then (b), then (d) of Proposition 4.5.1 give $\neg(\neg p \vee q)$, $\neg\neg p \wedge \neg q$, $p \wedge \neg q$.

(c) $\neg p \wedge q$: distributivity and commutativity give $(p \wedge \neg p) \vee (q \wedge \neg p)$; then (c) of Proposition 4.9.4 gives $\bot \vee (q \wedge \neg p)$, and (a) with commutativity gives $\neg p \wedge q$.

(d) p: (g) read from right to left gives $p \wedge (q \vee \neg q)$, then (c) and (a) of Proposition 4.9.4.

(e) Positive: the $\neg$ flips it, and so does the hypothesis place of $\rightarrow$.

(f) Negative: the right side of $\rightarrow$ keeps the sign, the $\neg$ flips it.

(g) Unsigned: it is a part of a biconditional.

(h) $(r \rightarrow (p \wedge q)) \models (r \rightarrow p)$: $p \wedge q$ sits in the conclusion, a positive place.

(i) $(p \wedge q) \vee (\neg p \wedge \neg q)$, one minterm per T-row.

(j) $(\neg p \vee q) \wedge (p \vee \neg q)$, one clause per F-row.

(k) 8: one for each of the eight rows, since all are T-rows.

(l) Yes: $\alpha \vee \beta \equiv \neg\alpha \rightarrow \beta$, so a formula in $\{\neg, \vee\}$, adequate by Theorem 4.9.6, can be rewritten into $\{\neg, \rightarrow\}$, from the leaves up.

4.9.9.2 B. Practice

Exercise. B1: Normal forms of three formulas

Find a DNF and a CNF of each formula, by either route, and check one of them in two rows.

(a) $p \leftrightarrow \neg q$.

(b) $(p \wedge q) \rightarrow r$.

(c) $\neg(p \vee (q \rightarrow r))$.

Solution. (a) By the table: $p \leftrightarrow \neg q$ is true when p and q have opposite values, so its T-rows are T, F and F, T, and its F-rows are T, T and F, F. DNF: $(p \wedge \neg q) \vee (\neg p \wedge q)$, as in Example (Exactly one of two). CNF, one clause per F-row from the opposite literals: the row T, T gives $\neg p \vee \neg q$ and the row F, F gives $p \vee q$, so $(\neg p \vee \neg q) \wedge (p \vee q)$. Check the CNF: in the row T, F both clauses are true, and so is $p \leftrightarrow \neg q$; in the row T, T the first clause is false, and so is $p \leftrightarrow \neg q$.

(b) By rewriting: $(p \wedge q) \rightarrow r \equiv \neg(p \wedge q) \vee r \equiv \neg p \vee \neg q \vee r$. This is a single clause, so it is a CNF, and it is also a DNF with three one-literal conjunctive clauses. Check: in the row T, T, F it is false (every literal false), as is $(p \wedge q) \rightarrow r$; in the row F, T, F it is true, as is the original (its hypothesis is false).

(c) By rewriting: $\neg(p \vee (q \rightarrow r)) \equiv \neg p \wedge \neg(q \rightarrow r) \equiv \neg p \wedge \neg(\neg q \vee r) \equiv \neg p \wedge q \wedge \neg r$. This single conjunctive clause is a DNF, and it is also a CNF with three one-literal clauses. Check: in the row F, T, F it is true, and so is the original ($q \rightarrow r$ is false and p is false); in the row F, T, T it is false, and so is the original ($q \rightarrow r$ is true).

Exercise. B2: Two connectives from two

Find formulas in $\{\neg, \wedge\}$, with atoms p and q, equivalent to $p \vee q$ and to $p \leftrightarrow q$. Check the second by a truth table.

Solution. $p \vee q \equiv \neg(\neg p \wedge \neg q)$: this is the translation t_1 in the proof of Theorem 4.9.6, and the right side is F exactly when $\neg p$ and $\neg q$ are both true, that is, when $p \vee q$ is F.

For the biconditional, $p \leftrightarrow q$ says $p \to q$ and $q \to p$, and $p \to q$ says "not p without q", that is, $\neg(p \wedge \neg q)$. So $p \leftrightarrow q \equiv \neg(p \wedge \neg q) \wedge \neg(\neg p \wedge q)$. The table:

p	q	$\neg(p \wedge \neg q)$	$\neg(\neg p \wedge q)$	$\neg(p \wedge \neg q) \wedge \neg(\neg p \wedge q)$	$p \leftrightarrow q$
T	T	T	T	T	T
T	F	F	T	F	F
F	T	T	F	F	F
F	F	T	T	T	T

The last two columns agree in every row.

Exercise. B3: Odd parity

Give an example of a formula in DNF, with atoms p, q, r, that is true in exactly the rows where an **odd** number of p, q, r are true. List the properties your answer must have, and check it against them.

Solution. The rows with an odd number of T's are those with three T's, T T T, and those with one, T F F, F T F, and F F T. One minterm per row, as in Theorem 4.9.1:

$$(p \wedge q \wedge r) \vee (p \wedge \neg q \wedge \neg r) \vee (\neg p \wedge q \wedge \neg r) \vee (\neg p \wedge \neg q \wedge r).$$

The properties an answer must have: (i) it is in DNF (a disjunction of conjunctions of literals); (ii) its atoms are among p, q, r; (iii) it is T in the four rows T T T, T F F, F T F, F F T; (iv) it is F in the other four, T T F, T F T, F T T, F F F. Check: (i) and (ii) hold by inspection. By Eq. 4.9.2, each minterm is T in its own row only, so the disjunction is T exactly in the four rows listed, which gives (iii) and (iv).

A different correct answer, for instance the same minterms in another order, or with the literals of a minterm reordered, passes the same checks. No DNF with fewer conjunctive clauses works. A conjunctive clause containing an atom and its negation is T in no row and contributes nothing. Any other conjunctive clause omitting an atom, say $p \wedge q$, is T in two rows that differ at that atom, and changing one atom always changes the parity, so one of those rows has even parity. So every conjunctive clause that is T somewhere mentions all three atoms and is T in exactly one row, and four T-rows need at least four conjunctive clauses.

4.9.9.3 C. Going deeper

Exercise. C1: One connective is enough

Add to the language two binary connectives: $\alpha \mid \beta$ ("not both", NAND), which is F exactly when α and β are both T, and $\alpha \downarrow \beta$ ("neither", NOR), which is T exactly when α and β are both F.

(a) Prove that $\{|\}$ is adequate.

(b) Prove that $\{\downarrow\}$ is adequate.

Solution. (a) First, two equivalences, checked from the table of $|$: $\alpha \mid \alpha$ is F exactly when α is T, so $\alpha \mid \alpha \equiv \neg\alpha$; and $(\alpha \mid \beta) \mid (\alpha \mid \beta) \equiv \neg(\alpha \mid \beta)$, which is T exactly when α and β are both T, so it is equivalent to $\alpha \wedge \beta$. Now define, by recursion on formulas as in Theorem 4.3.4, $s(p) := p$, $s(\neg\alpha) := s(\alpha) \mid s(\alpha)$, $s(\alpha \wedge \beta) := (s(\alpha) \mid s(\beta)) \mid (s(\alpha) \mid s(\beta))$ (the other clauses do not matter). By structural induction, exactly as in the Claim in the proof of Theorem 4.9.6, every formula φ in $\{\neg, \wedge\}$ has $s(\varphi)$ in $\{|\}$ and $v(s(\varphi)) = v(\varphi)$ for every v: the atom case is immediate, and the two equivalences above give the $\neg$ and $\wedge$ cases from the induction hypothesis. Given a truth function f, Theorem 4.9.6 (a) gives a formula φ in $\{\neg, \wedge\}$ expressing it, and then $s(\varphi)$ expresses f and is in $\{|\}$.

(b) The same, with $\{\neg, \vee\}$: $\alpha \downarrow \alpha$ is T exactly when α is F, so $\alpha \downarrow \alpha \equiv \neg\alpha$; and $(\alpha \downarrow \beta) \downarrow (\alpha \downarrow \beta) \equiv \neg(\alpha \downarrow \beta) \equiv \alpha \vee \beta$. The recursive translation $s(\neg\alpha) := s(\alpha) \downarrow s(\alpha)$, $s(\alpha \vee \beta) := (s(\alpha) \downarrow s(\beta)) \downarrow (s(\alpha) \downarrow s(\beta))$ turns a formula in $\{\neg, \vee\}$ expressing f into one in $\{\downarrow\}$.

(Recursion and structural induction for the language with the new connectives hold by the same proofs as Theorem 4.3.4 and Theorem 4.3.1, with one more formation clause.)

Marking scheme. (10 points)

- 3: (a) two equivalences expressing the connectives of an adequate set from Theorem 4.9.6 (a) in terms of $|$, read off its table (for instance $\alpha \mid \alpha \equiv \neg\alpha$ and $(\alpha \mid \beta) \mid (\alpha \mid \beta) \equiv \alpha \wedge \beta$).
- 3: (a) the recursive translation and the structural induction showing that it keeps the value under every valuation.
- 1: (a) the conclusion, via Theorem 4.9.6 (a).
- 3: (b) the same for $\downarrow$, with $\{\neg, \vee\}$.

Fatal errors (0 for the step): checking the translation on examples only, with no induction.

Exercise. C2: Two sets that are not adequate

(a) Prove that $\{\wedge, \vee, \rightarrow, \leftrightarrow\}$ is not adequate.

(b) Add to the language a three-place connective $\mathsf{M}(\alpha, \beta, \gamma)$, **majority**, which is T exactly when at least two of α, β, γ are T. Prove that $\{\neg, \mathsf{M}\}$ is not adequate.

Solution (Hint 1). For (b): compare each row with the row in which every atom's value is flipped.

Solution (Hint 2). For (b): show that every formula in $\neg$ and M takes opposite values in a row and in its flipped row.

Solution. *(a)* Let v_{T} make every atom true. Every formula in $\{\wedge, \vee, \to, \leftrightarrow\}$ is true under v_{T}, by the structural induction in the proof of Theorem 4.9.6 (b) with one more case: if $v_{\mathrm{T}}(\alpha) = v_{\mathrm{T}}(\beta) = \mathrm{T}$, then $v_{\mathrm{T}}(\alpha \leftrightarrow \beta) = \mathrm{T}$, since the two values are equal. As there, no such formula expresses negation, which must give F at the input T.

(b) For a valuation v, let $\bar{v}$ be the **flipped** valuation: $\bar{v}(p)$ is the opposite of $v(p)$, for every atom p. We show by structural induction that every formula φ in $\{\neg, \mathsf{M}\}$ satisfies: for every valuation v, $\bar{v}(\varphi)$ is the opposite of $v(\varphi)$. For an atom this is the definition of $\bar{v}$. If $\bar{v}(\alpha)$ is the opposite of $v(\alpha)$, then $\bar{v}(\neg\alpha)$, the opposite of $\bar{v}(\alpha)$, is $v(\alpha)$, which is the opposite of $v(\neg\alpha)$. If the property holds for α, β, γ, then flipping v flips all three values; at least two of them are T after the flip exactly when at least two were F before, that is, exactly when at most one was T, that is, exactly when $\mathsf{M}(\alpha, \beta, \gamma)$ was F. So $\bar{v}(\mathsf{M}(\alpha, \beta, \gamma))$ is the opposite of $v(\mathsf{M}(\alpha, \beta, \gamma))$. The other connectives and $\bot$ do not occur in formulas of $\{\neg, \mathsf{M}\}$, so those cases are vacuous. (Structural induction for the extended language holds by the proof of Theorem 4.3.1, with one more formation clause.)

Now the two-input function of $p_1 \wedge p_2$ is F at (T, F) and at the flipped input (F, T). If a formula φ in $\{\neg, \mathsf{M}\}$ expressed it, then for a valuation v with $v(p_1) = \mathrm{T}$, $v(p_2) = \mathrm{F}$ we would have $v(\varphi) = \mathrm{F}$ and $\bar{v}(\varphi) = \mathrm{F}$, since $\bar{v}(p_1) = \mathrm{F}$ and $\bar{v}(p_2) = \mathrm{T}$. But these values must be opposite. So no formula in $\{\neg, \mathsf{M}\}$ expresses it, and the set is not adequate.

Marking scheme. (10 points)

- 3: (a) states the all-T valuation and runs the induction, with the $\leftrightarrow$ case and the vacuous cases ($\neg$, $\bot$) named.
- 1: (a) names a truth function that is not expressed (negation), with the reason.
- 4: (b) states the flipped-row property for **every** formula in $\{\neg, \mathsf{M}\}$ and proves it by structural induction, including the M case.
- 2: (b) exhibits a truth function that takes the same value on some row and on its flipped row, and concludes.

Fatal errors (0 for the step): in (b), checking the flipped-row property on a few formulas instead of proving it for all; in (a), an implication case justified by the hypothesis being true rather than by the conclusion.

Exercise. C3: Exit check: normal forms

Put $(p \to q) \to (q \wedge r)$ into DNF and into CNF, by either route, and check one of your answers against the truth table.

Solution. *By rewriting.* Eliminate the outer implication and then the inner one, and push the negation in:

$$\begin{aligned}(p \to q) \to (q \wedge r) &\equiv \neg(p \to q) \vee (q \wedge r)\\ &\equiv \neg(\neg p \vee q) \vee (q \wedge r)\\ &\equiv (p \wedge \neg q) \vee (q \wedge r).\end{aligned}$$

The last line is a DNF. Here $\neg(p \to q) \equiv p \wedge \neg q$, since an implication is false exactly when its hypothesis is true and its conclusion false. For a CNF, distribute:

$$(p \wedge \neg q) \vee (q \wedge r) \equiv (p \vee q) \wedge (p \vee r) \wedge (\neg q \vee q) \wedge (\neg q \vee r),$$

and drop the clause $\neg q \vee q$, which is true in every row: the CNF is $(p \vee q) \wedge (p \vee r) \wedge (\neg q \vee r)$. (The clause $p \vee r$ is redundant too: $(p \vee q) \wedge (\neg q \vee r)$ is an equivalent CNF, as splitting on q shows. Either is correct.)

The check. The formula is false exactly when $p \to q$ holds and $q \wedge r$ fails. Its table:

p	q	r	$p \to q$	$q \wedge r$	$(p \to q) \to (q \wedge r)$
T	T	T	T	T	T
T	T	F	T	F	F
T	F	T	F	F	T
T	F	F	F	F	T
F	T	T	T	T	T
F	T	F	T	F	F
F	F	T	T	F	F
F	F	F	T	F	F

The T-rows are T T T, T F T, T F F, and F T T. The DNF $(p \wedge \neg q) \vee (q \wedge r)$ is true in exactly these: $p \wedge \neg q$ in T F T and T F F, and $q \wedge r$ in T T T and F T T. In each of the four F-rows the CNF has a false clause: $\neg q \vee r$ in T T F and F T F, $p \vee q$ in F F T and F F F.

By the table route, the DNF has four minterms, one per T-row, and the CNF four clauses, one per F-row, for instance $\neg p \vee \neg q \vee r$ for the row T T F. Any correct DNF and CNF earns full marks; a CNF is not unique.

Marking scheme. (10 points)

- 2: the route carried out correctly: $\to$ eliminated, with $\neg(p \to q) \equiv p \wedge \neg q$ and its reason, or the T-rows and F-rows read off a correct table.
- 3: a correct DNF.
- 3: a correct CNF, by distribution carried out in full or by one clause per F-row.
- 2: the check against the table, covering at least the four F-rows (or the four T-rows).

Fatal errors (0 for the step): $\neg(p \to q)$ rewritten as $\neg p \to \neg q$; a "CNF" with a conjunction inside a clause.

Exercise. C4: From connectives to sets

Let U be a collection, and A, B, E collections of members of U, in the everyday sense; $U \setminus A$ is the collection of members of U not in A. For a member x of U, write a, b, e for the statements $x \in A$, $x \in B$, $x \in E$.

(a) Write "$x \in U \setminus (A \cap B)$" and "$x \in (U \setminus A) \cup (U \setminus B)$" as formulas in a, b, and deduce from a standard equivalence that the two collections have the same members.

(b) Show in the same way that $A \cap (B \cup E)$ and $(A \cap B) \cup (A \cap E)$ have the same members.

(c) Show that $(a \to b) \to ((a \wedge e) \to (b \wedge e))$ is a tautology, and deduce: if every member of A is in B, then every member of $A \cap E$ is in $B \cap E$.

Solution. *(a)* The formulas are $\neg(a \wedge b)$ and $\neg a \vee \neg b$. They are equivalent by Proposition 4.5.1 (b). Each member x of U gives a valuation of a and b, by whether it lies in A and in B, and under that valuation the two formulas have the same value. So x is in one collection exactly when it is in the other.

(b) The formulas are $a \wedge (b \vee e)$ and $(a \wedge b) \vee (a \wedge e)$, equivalent by Proposition 4.5.1 (g); conclude as in (a).

(c) Work backwards from F. Falsity needs $a \to b$ true and $(a \wedge e) \to (b \wedge e)$ false, so a and e true and $b \wedge e$ false. Then b is true, by $a \to b$, so $b \wedge e$ is true: no row makes the formula false, and it is a tautology. For each x, its memberships give a valuation making the formula true: so if "$x \in A$ implies $x \in B$" holds, then "$x \in A \cap E$ implies $x \in B \cap E$" holds. If every member of A is in B, the first holds for every x, hence so does the second.

Marking scheme. (8 points)

- 3: (a), with the step from an equivalence to equal membership argued (each x gives a valuation).
- 2: (b).
- 3: (c): the tautology, and the step to collections.

Fatal errors (0 for the step): checking the set equality on one example collection only.

Exercise. C5: Counting rows

For a formula φ whose atoms are among $p_1, \ldots, p_n$, let $\#\varphi$ be the number of rows of the table on $p_1, \ldots, p_n$ in which φ is true.

(a) Show that $\varphi \models \psi$ implies $\#\varphi \leq \#\psi$, and that the converse fails.

(b) Show that $\#(\varphi \vee \psi) = \#\varphi + \#\psi - \#(\varphi \wedge \psi)$.

(c) Show that the minterms of two different rows are never true together, and that the disjunction of all 2^n minterms is a tautology. Deduce that, for satisfiable φ, the full DNF of φ over $p_1, \ldots, p_n$ (one minterm on $p_1, \ldots, p_n$ for each row on them in which φ is true) has exactly $\#\varphi$ minterms.

Solution. *(a)* If $\varphi \models \psi$, every row in which φ is true is one in which ψ is true, so there are at least as many of the second. The converse fails: on the atoms p, q, the formulas p and q are each true in two rows, so $\#p \leq \#q$, but $p \not\models q$ (the row $p = \mathrm{T}$, $q = \mathrm{F}$).

(b) By (V4), the rows of $\varphi \vee \psi$ are those of φ together with those of ψ. By (V3), the rows in both are the rows of $\varphi \wedge \psi$. The sum $\#\varphi + \#\psi$ counts those twice and every other row of $\varphi \vee \psi$ once.

(c) Two different rows differ at some p_i; one minterm contains p_i and the other $\neg p_i$, and no valuation makes both literals true. Each row makes its own minterm true (Eq. 4.9.2), so the disjunction of all 2^n is true in every row (Eq. 4.9.1). The full DNF over $p_1, \ldots, p_n$ has one minterm for each row on them in which φ is true, and different rows give different minterms, so it has $\#\varphi$ of them. (It is not always the DNF that the proof of Theorem 4.9.1 builds. That proof uses only the atoms of φ, so when φ is satisfiable and has at least one atom, and j of the p_i do not occur in φ, it gives $\#\varphi/2^j$ minterms on the remaining atoms. When φ has no atoms it gives $p \vee \neg p$ or $p \wedge \neg p$. When $\#\varphi = 0$ it gives a contradiction $q \wedge \neg q$ for some

atom q, which is a conjunctive clause but no minterm. Hence the restriction to satisfiable φ.)

Marking scheme. (8 points)

- 2: (a) the inequality; 1: a counterexample to the converse, checked.
- 2: (b), with the rows counted twice named.
- 3: (c): different rows give minterms never true together; the disjunction of all is a tautology; the count.

Fatal errors (0 for the step): in (a), a "converse" argued from one pair of formulas that happen to entail each other.

Exercise. C6: Challenge: Interpolation

(a) For $\varphi = p \wedge q$ and $\psi = q \vee r$, find a formula θ whose only atom is q, with $\varphi \models \theta$ and $\theta \models \psi$.

(b) Prove: if $\varphi \models \psi$, φ is satisfiable and ψ is not a tautology, then there is a formula θ, all of whose atoms occur both in φ and in ψ, with $\varphi \models \theta$ and $\theta \models \psi$.

(c) Show that under the hypotheses of (b), φ and ψ have at least one atom in common.

Solution (Hint 1). Let θ say: the shared atoms take values that some row making φ true also gives them.

Solution (Hint 2). To show $\theta \models \psi$, build one row that agrees with a row of φ on the atoms of φ and with the given row on the other atoms of ψ.

Solution. *(a)* $\theta = q$: $p \wedge q \models q$ and $q \models q \vee r$.

(c) (Proved first, since (b) uses it.) Suppose no atom is shared. Take a row u making φ true, and any row w. The row that agrees with u on the atoms of φ and with w on every other atom makes φ true (Lemma 4.4.1), hence ψ; it agrees with w on the atoms of ψ, none of which is an atom of φ, so w makes ψ true (Lemma 4.4.1 again). As w was any row, ψ would be a tautology.

(b) Let $q_1, \ldots, q_k$ be the shared atoms ($k \geq 1$ by (c)). Call a row s on $q_1, \ldots, q_k$ *good* if some valuation making φ true gives the q_i the values of s; there is at least one, since φ is satisfiable. Let θ be the disjunction of the minterms, on $q_1, \ldots, q_k$, of the good rows. Its atoms are shared.

$\varphi \models \theta$: a valuation making φ true gives the q_i the values of a good row s, so it makes the minterm of s true (Eq. 4.9.2), hence θ.

$\theta \models \psi$: let w make θ true. Then w makes the minterm of some good row s true, so w gives the q_i the values of s; and some u making φ true does the same. Let u' agree with u on the atoms of φ and with w on every other atom. By Lemma 4.4.1, u' makes φ true, so it makes ψ true, since $\varphi \models \psi$. Now u' agrees with w on every atom of ψ: on the shared atoms both

give the values of s, and the other atoms of ψ are not atoms of φ. By Lemma 4.4.1, w makes ψ true.

Marking scheme. (12 points)

- 1: (a).
- 2: the definition of θ from the good rows, with its existence (satisfiable φ).
- 2: $\varphi \models \theta$.
- 5: $\theta \models \psi$, with the glued row and both uses of Lemma 4.4.1.
- 2: (c).

Fatal errors (0 for the step): a θ with an atom of φ alone; a glued row that is not shown to agree with w on every atom of ψ.

4.10 Compactness and Maximal Consistent Sets

This section is optional, and nothing outside it depends on it.

Every list of premises in this chapter has been finite, and the proof of completeness needed that: its Step 1 lists the atoms of Γ and ψ, and there are only finitely many of them because Γ is a finite list (Theorem 4.8.4). But a statement about infinitely many objects can come with infinitely many premises. Give each $n \geq 1$ an atom p_n, read "the n-th lamp is on", and the rules "the first lamp is on" and "if the n-th lamp is on, so is the next" are one premise for each n. This section defines $\vdash$ and $\models$ for such collections of premises, proves completeness for them by a route different from that of Section 4.8, and deduces from it the compactness theorem: a collection of formulas can be made true all at once as soon as each finite part of it can. The word "collection" keeps its everyday sense throughout; nothing here needs more.

4.10.1 Infinitely many premises

A derivation is finite, so from infinitely many premises it can use only finitely many; a valuation, on the other hand, gives every formula a value at once.

Definition 4.10.1. Derivability and Entailment from Infinitely Many Premises

Let Γ be a collection of formulas, finite or infinite, and ψ a formula. $\Gamma \vdash \psi$ if ψ is derivable from some finite list of members of Γ; Γ is **consistent** if $\Gamma \nvdash \bot$. $\Gamma \models \psi$ if every valuation that makes every member of Γ true makes ψ true; Γ is **satisfiable** if some valuation makes every member of Γ true.

In words: a derivation has finitely many lines, so it can cite only finitely many premises. That is why $\vdash$ goes through finite lists of members, while $\models$ looks at all of Γ at once. For a finite list the new definitions agree with the old. A derivation from part of a list is a derivation from the list, by Lemma 2.3.3 (a), so $\vdash$ means what Definition 1.2.3 said, and *consistent* means what Definition 3.2.1 said. The definitions of $\models$ and *satisfiable* say for a finite list exactly what Definition 4.5.3 and Definition 4.5.2 said. "Finite list of members" allows repetitions and any order; neither matters, by Lemma 2.3.3 (a) and by the definitions of $\models$ and *satisfiable*.

An example. Let Γ consist of p_1 and every formula $p_n \to p_{n+1}$ with $n \geq 1$. Then $\Gamma \vdash p_3$:

from the finite list p_1, $p_1 \to p_2$, $p_2 \to p_3$, two $\to$E lines give p_2 and then p_3. And $\Gamma \models p_n$ for every $n \geq 1$: a valuation making every member of Γ true makes p_1 true, and makes p_{n+1} true whenever it makes p_n true, so it makes every p_n true, by Axiom 4.2.1 on n, started at 1.

Quick check. Is the collection consisting of $p_1 \vee p_2$ and every $\neg p_n$ with $n \geq 2$ satisfiable? Is it consistent?

Solution. Satisfiable: the valuation making p_1 true and every other atom false makes every member true. Hence consistent. A derivation of $\bot$ would use some finite list of members; that list is made true by the same valuation, so it is satisfiable, hence consistent by Section 4.7, and derives no $\bot$.

4.10.2 Listing all formulas

The construction below goes through all formulas one at a time, so it needs them in one list.

Lemma 4.10.1. The Formulas Can Be Listed

There is a list $\varphi_0, \varphi_1, \varphi_2, \ldots$ of formulas in which every formula occurs.

In words: the atoms are listed already $(p, q, r, p_1, p_2, \ldots)$, and each formula is a finite string of symbols, so the formulas can be listed by size, finitely many at each stage.

↻ **Second reading** · *How to list every formula*

Idea. List by size. At stage n, allow only the first n atoms and strings of at most n symbols. Each stage is finite, and every formula fits into some stage.

Proof. Write the atoms in one list $a_1, a_2, a_3, \ldots$ (say $p, q, r, p_1, p_2, \ldots$ in this order). For each $n \geq 1$, consider the strings of length at most n over the $n + 8$ symbols

$$a_1, \ldots, a_n,\ \bot,\ \neg,\ \wedge,\ \vee,\ \to,\ \leftrightarrow,\ (,\ ).$$

There are finitely many such strings (at most $(n+8)^n$ of each length up to n), so finitely many of them are formulas. Stage n lists, in any order, the formulas among them that no earlier stage listed. Write the stages one after another and number the entries $\varphi_0, \varphi_1, \ldots$. Each stage n lists at least the atom a_n, which is available for the first time, so the stages together list infinitely many formulas.

Every formula occurs. Let φ have length ℓ, and let its atoms be among $a_1, \ldots, a_m$ (with $m = 0$ if it has none). At stage $n = \max\{\ell, m\}$ (note $\ell \geq 1$), φ is a string of length at most n over the symbols of that stage, so it is listed then, or earlier.

■

4.10.3 Lindenbaum's construction

The new proof of completeness works the other way round from Section 4.8. It does not build a derivation. It starts from a consistent collection and builds a row that makes all of it true. The row is read off a collection that has made up its mind about every formula.

A collection Δ is **maximal consistent** if it is consistent and, for every formula φ, φ is in Δ or $\neg\varphi$ is.

The construction rests on one step. Exercise C2 of Section 3.2 (Exercise (C2: Challenge: Splitting an inconsistency, used later)) proved it for finite lists; here it is again, for collections, with its proof.

The key step. *If Δ is consistent, then Δ with φ added, or Δ with $\neg\varphi$ added, is consistent.* Suppose both were inconsistent. Then there are finite lists Δ_1 and Δ_2 of members of Δ with $\Delta_1, \varphi \vdash \bot$ and $\Delta_2, \neg\varphi \vdash \bot$ (if a derivation of $\bot$ does not use φ, or $\neg\varphi$, adding it as a premise changes nothing, by weakening). By weakening, both hold with the list Δ_1, Δ_2 in place of Δ_1 and of Δ_2. A $\neg$I box on φ gives $\Delta_1, \Delta_2 \vdash \neg\varphi$, and an RAA box on $\neg\varphi$ gives $\Delta_1, \Delta_2 \vdash \varphi$ (this is Proposition 3.2.2). Then $\neg$E, by chaining (Lemma 2.3.3 (c)), gives $\Delta_1, \Delta_2 \vdash \bot$, and Δ would be inconsistent.

Lemma 4.10.2. Lindenbaum's Lemma

Every consistent collection of formulas is contained in a maximal consistent collection.

Idea. Go through the list of all formulas, and add each one or its negation, whichever keeps the collection consistent. The key step says one of them always does. At the end, every formula has been decided, and no contradiction has been let in, because a contradiction would show up at some finite stage.

Proof. Let Γ be consistent, and let $\varphi_0, \varphi_1, \ldots$ be the list of Lemma 4.10.1. Define collections $\Delta_0, \Delta_1, \ldots$ by recursion on n, each from the one before:

$$\Delta_0 := \Gamma; \qquad \Delta_{n+1} := \begin{cases} \Delta_n \text{ with } \varphi_n \text{ added}, & \text{if that collection is consistent;} \\ \Delta_n \text{ with } \neg\varphi_n \text{ added}, & \text{otherwise.} \end{cases}$$

Every Δ_n is consistent, by Axiom 4.2.1: $\Delta_0 = \Gamma$ is; and if Δ_n is, then Δ_{n+1} is, by its definition in the first case and by the key step in the second. Each Δ_n contains the one before.

Let Δ consist of the formulas that belong to some Δ_n. Then Δ contains $\Gamma = \Delta_0$.

Δ is consistent. A derivation of $\bot$ from Δ uses a finite list of members of Δ. Each of them is in some Δ_n, so all of them are in the Δ_N with the largest of those indices N ($N = 0$ if the list is empty), since the collections grow. That derivation would then show $\Delta_N \vdash \bot$, but Δ_N is consistent.

Δ is maximal. Every formula is some φ_n, and φ_n or $\neg\varphi_n$ is in Δ_{n+1}, hence in Δ. This proves the lemma.

■

Quick check. Start from $\Gamma = \{p \vee q,\ \neg p\}$ with a list of formulas beginning p, q. What are Δ_1 and Δ_2?

Solution. Γ with p added is inconsistent (p and $\neg p$ give $\bot$ by $\neg$E), so Δ_1 adds $\neg p$, which is already there: $\Delta_1 = \Gamma$. Δ_1 with q added is satisfiable ($p = \mathrm{F}$, $q = \mathrm{T}$), hence consistent (Section 4.7), so Δ_2 adds q: $\Delta_2 = \{p \vee q,\ \neg p,\ q\}$.

4.10.4 A row from a maximal consistent collection

A maximal consistent collection behaves like the collection of formulas that are true in one row: it contains a negation exactly when it lacks the negated formula, a conjunction exactly when it contains both parts, and so on through the clauses of Definition 4.4.2.

Lemma 4.10.3. Maximal Consistent Collections

Let Δ be maximal consistent, and let φ, ψ be formulas.

(a) If $\Delta \vdash \varphi$, then $\varphi \in \Delta$.

(b) Exactly one of φ, $\neg\varphi$ is in Δ; and $\bot \notin \Delta$.

(c) $\varphi \wedge \psi \in \Delta$ iff $\varphi \in \Delta$ and $\psi \in \Delta$; $\varphi \vee \psi \in \Delta$ iff $\varphi \in \Delta$ or $\psi \in \Delta$; $\varphi \to \psi \in \Delta$ iff $\varphi \notin \Delta$ or $\psi \in \Delta$; $\varphi \leftrightarrow \psi \in \Delta$ iff φ and ψ are both in Δ or both not in Δ.

In words: whatever Δ derives, it already contains; it decides every formula, one way only; and it treats each connective as the truth table does.

↻ **Second reading** · *The proof: a maximal collection decides everything*

Idea. For each clause, find the derivable sequent that says the same thing as the truth table, and let (a) put what Δ derives into Δ; where a formula must be shown missing, derive its negation instead and use (b).

Proof. Derivations from Δ are combined by Lemma 2.3.3 throughout: two finite lists of members of Δ can be joined into one, and a sequent already proved can be applied by chaining.

(a) Suppose $\Delta \vdash \varphi$ and $\varphi \notin \Delta$. By maximality $\neg\varphi \in \Delta$, so $\Delta \vdash \neg\varphi$ by a one-line derivation, and $\neg$E gives $\Delta \vdash \bot$, against consistency.

(b) At least one of φ, $\neg\varphi$ is in Δ, by maximality. If both were, $\neg$E would give $\Delta \vdash \bot$. And $\bot \in \Delta$ would give $\Delta \vdash \bot$ at once.

(c) Each clause comes from (a), (b) and a derivable sequent.

$\wedge$. If $\varphi \wedge \psi \in \Delta$, then $\wedge$E gives $\Delta \vdash \varphi$ and $\Delta \vdash \psi$, so both are in Δ by (a). If both are in Δ, $\wedge$I gives $\Delta \vdash \varphi \wedge \psi$, so it is in Δ by (a).

$\vee$. If $\varphi \in \Delta$ or $\psi \in \Delta$, then $\vee$I gives $\Delta \vdash \varphi \vee \psi$, so it is in Δ by (a). If neither is in Δ, then $\neg\varphi, \neg\psi \in \Delta$ by (b), and ($\vee$3) of Lemma 4.8.1 gives $\Delta \vdash \neg(\varphi \vee \psi)$; so $\neg(\varphi \vee \psi) \in \Delta$ by (a), and $\varphi \vee \psi \notin \Delta$ by (b).

$\rightarrow$. If $\varphi \rightarrow \psi \in \Delta$ and $\varphi \in \Delta$, then $\rightarrow$E gives $\Delta \vdash \psi$, so $\psi \in \Delta$ by (a); hence $\varphi \notin \Delta$ or $\psi \in \Delta$. Conversely, if $\varphi \notin \Delta$, then $\neg\varphi \in \Delta$ by (b), and ($\rightarrow$1) gives $\Delta \vdash \varphi \rightarrow \psi$; if $\psi \in \Delta$, then ($\rightarrow$2) gives it. Either way $\varphi \rightarrow \psi \in \Delta$ by (a).

$\leftrightarrow$. If both φ, ψ are in Δ, ($\leftrightarrow$1) gives $\Delta \vdash \varphi \leftrightarrow \psi$; if neither, then $\neg\varphi, \neg\psi \in \Delta$ and ($\leftrightarrow$2) gives it; in both cases $\varphi \leftrightarrow \psi \in \Delta$ by (a). If exactly one is in Δ, the other's negation is in Δ by (b), and ($\leftrightarrow$3) or ($\leftrightarrow$4) gives $\Delta \vdash \neg(\varphi \leftrightarrow \psi)$; so $\neg(\varphi \leftrightarrow \psi) \in \Delta$ by (a), and $\varphi \leftrightarrow \psi \notin \Delta$ by (b). (With $\varphi \leftrightarrow \psi \in \Delta$, $\leftrightarrow$E passes membership from either part to the other, which is the same fact read from the other side.)

■

So a maximal consistent collection names a row: make an atom true exactly when the collection contains it.

Lemma 4.10.4. Truth Lemma

Let Δ be maximal consistent, and let v_Δ be the valuation that makes an atom p true exactly when $p \in \Delta$. Then for every formula φ: $v_\Delta(\varphi) = \mathrm{T}$ iff $\varphi \in \Delta$.

Idea. Induct on φ. Each clause of Definition 4.4.2 has a twin in Lemma 4.10.3: part (b) for $\neg$, part (c) for the binary connectives. The induction hypothesis turns "true under v_Δ" into "in Δ" for the parts, and the twin does the rest.

Proof. By Theorem 4.3.1, with the property "$v_\Delta(\varphi) = \mathrm{T}$ iff $\varphi \in \Delta$".

Atoms: this is the definition of v_Δ. $\bot$*:* $v_\Delta(\bot) = \mathrm{F}$ by (V1), and $\bot \notin \Delta$ by Lemma 4.10.3 (b), so both sides of the "iff" fail.

$\neg\varphi$*:* $v_\Delta(\neg\varphi) = \mathrm{T}$ iff $v_\Delta(\varphi) = \mathrm{F}$, by (V2); iff $\varphi \notin \Delta$, by the induction hypothesis; iff $\neg\varphi \in \Delta$, by Lemma 4.10.3 (b).

$\varphi \wedge \psi$*:* $v_\Delta(\varphi \wedge \psi) = \mathrm{T}$ iff $v_\Delta(\varphi) = \mathrm{T}$ and $v_\Delta(\psi) = \mathrm{T}$, by (V3); iff $\varphi \in \Delta$ and $\psi \in \Delta$, by the induction hypothesis; iff $\varphi \wedge \psi \in \Delta$, by Lemma 4.10.3 (c). The cases $\vee$, $\rightarrow$, $\leftrightarrow$ are the same: the clause of Definition 4.4.2 for the connective and the clause of Lemma 4.10.3 (c) for it say the same thing, once "true under v_Δ" is read as "in Δ" for the two parts, as the induction hypothesis allows. This proves the lemma.

■

4.10.5 Completeness again

Theorem 4.10.5. Completeness for Infinitely Many Premises

For every collection Γ of formulas, finite or infinite, and every formula ψ: if $\Gamma \models \psi$, then $\Gamma \vdash \psi$. Equivalently, every consistent collection is satisfiable.

Idea. If Γ does not derive ψ, then adding $\neg\psi$ cannot lead to $\bot$, since an RAA box would turn such a derivation into one of ψ. Extend Γ with $\neg\psi$ to a maximal consistent collection and read off its row: it makes every member of Γ true and ψ false.

Proof. We prove the contrapositive of the first sentence. Suppose $\Gamma \nvdash \psi$. Then Γ with $\neg\psi$ added is consistent. For suppose some finite list of its members derived $\bot$; adding $\neg\psi$ to the list if it is not there (weakening), we have $\Gamma_0, \neg\psi \vdash \bot$ for a finite list Γ_0 of members of Γ, and an RAA box on $\neg\psi$ gives $\Gamma_0 \vdash \psi$, so $\Gamma \vdash \psi$. By Lemma 4.10.2, Γ with $\neg\psi$ is contained in a maximal consistent Δ. By Lemma 4.10.4, v_Δ makes every member of Δ true, in particular every member of Γ, and $\neg\psi$. So v_Δ makes ψ false, and $\Gamma \not\models \psi$.

The equivalence: by Definition 4.10.1 and (V1), Γ is satisfiable iff $\Gamma \not\models \bot$, and consistent iff $\Gamma \nvdash \bot$. So the first sentence with $\psi = \bot$ is the second. Conversely, given the second, if $\Gamma \nvdash \psi$, then Γ with $\neg\psi$ is consistent, as just shown, hence satisfiable, by a row that refutes $\Gamma \models \psi$.

■

Two proofs, two prices. Section 4.8 glued the rows of a truth table into one derivation. That is a construction one can carry out, from 2^n row derivations for n atoms, but only for finitely many premises. This section never writes a derivation. It shows that if none exists, then a row exists, by a construction that passes through infinitely many formulas, one at a time, and cannot be finished by hand. The first proof is effective (it writes the derivation down) and limited to finite lists; the second reaches infinite collections and produces no derivation. The second is the route that extends beyond propositional logic.

Second reading · *A remark on the list of formulas*

Besides the rules, and the structural induction and recursion of Section 4.3, the proof used one new fact about the formulas: that they can be listed (Lemma 4.10.1). The construction then decides the formulas in the order of the list, and a different list may give a different Δ and a different row; any of them refutes $\Gamma \models \psi$. The proof cites Proposition 2.5.3 nowhere, but its key step uses RAA, and so does the step from $\Gamma \nvdash \psi$ to the consistency of Γ with $\neg\psi$. It is classical in the same way as the splitting of Section 4.8, which used excluded middle.

4.10.6 Compactness

Theorem 4.10.6. Compactness

A collection Γ of formulas is satisfiable if and only if every finite list of members of Γ is satisfiable. Consequently, $\Gamma \models \psi$ if and only if some finite list of members of Γ entails ψ.

Proof. "Only if": a valuation making every member of Γ true makes every member of each finite list true. "If": every finite list of members of Γ is satisfiable, hence consistent, by Section 4.7. So Γ is consistent, since a derivation of $\bot$ from Γ would use a finite list of members. By Theorem 4.10.5, Γ is satisfiable.

For the second sentence: $\Gamma \models \psi$ iff Γ with $\neg\psi$ added is unsatisfiable (a valuation making Γ true and ψ false is exactly one making both Γ and $\neg\psi$ true); iff, by the first sentence, some finite list of its members is unsatisfiable; iff some finite list Γ_0 of members of Γ has $\Gamma_0, \neg\psi$ unsatisfiable (adding $\neg\psi$ to an unsatisfiable list keeps it unsatisfiable, and $\Gamma_0, \neg\psi$ is itself such a list); iff $\Gamma_0 \models \psi$, by Section 4.5's reduction of entailment to unsatisfiability.

■

Example. Infinitely many premises, finitely many used

Let Γ consist of p_1 and every $p_n \to p_{n+1}$ with $n \geq 1$, the lamps of the opening.

(a) Show $\Gamma \models p_{100}$, and find a finite list of members of Γ that entails p_{100}.

(b) Show that a finite list of members of Γ that leaves out $p_{50} \to p_{51}$ does not entail p_{100}.

Solution. *(a)* A valuation making every member of Γ true makes p_1 true and passes truth from each p_n to p_{n+1}, so it makes every p_n true (Axiom 4.2.1, started at 1), p_{100} among them. The hundred formulas p_1, $p_1 \to p_2$, …, $p_{99} \to p_{100}$ already entail p_{100}: a row making them all true makes $p_1, p_2, \ldots, p_{100}$ true, one after another.

(b) Take the valuation making $p_1, \ldots, p_{50}$ true and every other atom false. It makes p_1 true; it makes $p_n \to p_{n+1}$ true for $n < 50$ (true conclusion) and for $n > 50$ (false hypothesis); only $p_{50} \to p_{51}$ is false. So it makes every member of the finite list true, and p_{100} false.

Compactness says that this is no accident: an entailment from infinitely many premises always rests on finitely many of them. Here every finite list that entails p_{100} must contain all ninety-nine implications up to $p_{99} \to p_{100}$, since (b)'s argument works with any $p_k \to p_{k+1}$, $k < 100$, in place of $p_{50} \to p_{51}$, and it must contain p_1, since the valuation making every atom false makes all the implications true.

Warning. Compactness promises **one** valuation for the whole collection when each finite part has one. The valuations of the finite parts may all differ, and the theorem does not say how to find the single one: its proof went through a maximal consistent collection, which nobody writes down.

Remark (Looking ahead). Chapter 7 states completeness and compactness for predicate logic (Section 7.6, optional), without proof; its sketch of completeness follows the route of this section.

4.10.7 Exercises

4.10.7.1 A. Check your understanding

Exercise. A1

(a) Why is $\Gamma \vdash \psi$ defined through finite lists of members of Γ?

(b) True or false: an infinite collection Γ is consistent iff every finite list of its members is consistent.

(c) In the proof of Lemma 4.10.2, why can Δ_n with φ_n added and Δ_n with $\neg\varphi_n$ added not both be inconsistent?

(d) Δ is maximal consistent, $p \in \Delta$ and $q \notin \Delta$. Is $p \to q$ in Δ? Is $\neg(p \wedge \neg q)$?

(e) Which result of Section 4.7 does the proof of Theorem 4.10.6 use?

Solution. *(a)* A derivation has finitely many lines, so it uses finitely many premises.

(b) True, by Definition 4.10.1: $\Gamma \vdash \bot$ means that some finite list of members derives $\bot$.

(c) By the key step: $\neg$I and RAA would derive $\neg\varphi_n$ and φ_n from finitely many members of Δ_n, and $\neg$E then $\bot$, making Δ_n inconsistent.

(d) Neither. $p \to q \notin \Delta$ by Lemma 4.10.3 (c), since $p \in \Delta$ and $q \notin \Delta$. And $\neg q \in \Delta$ by (b), so $p \wedge \neg q \in \Delta$ by (c), so $\neg(p \wedge \neg q) \notin \Delta$ by (b).

(e) That a satisfiable list is consistent (soundness).

Exercise. A2: Drill

The parts run from judging collections, through the definitions and single stages of the construction, to the lemmas and the theorem, and end with one more collection to judge.

(a) Is the collection of all p_n with $n \geq 1$ satisfiable?

(b) Is the collection of p_1 and every $\neg p_n$ with $n \geq 2$ consistent?

(c) Does the collection of all $p_n \to p_{n+1}$ with $n \geq 1$ entail $p_1 \to p_7$? If so, which finite list of members does?

(d) Is the collection of $p_1 \vee p_2$, $\neg p_1$, $\neg p_2$ and every p_n with $n \geq 3$ satisfiable?

(e) For a finite Γ, does Definition 4.10.1 change what $\Gamma \vdash \psi$ means?

(f) Run Lindenbaum's construction from $\Gamma = \{p\}$ with a list beginning $q, \neg p$. What are Δ_1 and Δ_2?

(g) Δ is maximal consistent and $p \notin \Delta$. Is $\neg p \in \Delta$? When is $p \vee q \in \Delta$?

(h) $p \in \Delta$ and $\neg q \in \Delta$, with Δ maximal consistent. Is $p \wedge \neg q \in \Delta$?

(i) Can a maximal consistent Δ contain $p \to q$, p and $\neg q$?

(j) How many formulas of length 1 contain no atom other than p and q?

(k) Can $\Gamma \models \psi$ hold with no finite list of members of Γ entailing ψ?

(l) Is the collection of all $p_n \leftrightarrow \neg p_{n+1}$ with $n \geq 1$ satisfiable?

Solution.

(a) Yes: every atom true.

(b) Yes: it is satisfiable (p_1 true, every other atom false), hence consistent.

(c) Yes; the list $p_1 \to p_2,\ \ldots,\ p_6 \to p_7$ does.

(d) No: its first three members already have no common row.

(e) No (Lemma 2.3.3 (a)).

(f) $\Delta_1 = \{p, q\}$, which is satisfiable, hence consistent. Δ_1 with $\neg p$ is inconsistent, so $\Delta_2 = \{p, q, \neg\neg p\}$.

(g) Yes, by Lemma 4.10.3 (b); and $p \vee q \in \Delta$ exactly when $q \in \Delta$, by (c).

(h) Yes, by Lemma 4.10.3 (c).

(i) No: $\to$E and $\neg$E would derive $\bot$.

(j) Three: p, q, $\bot$.

(k) No, by Theorem 4.10.6.

(l) Yes: make p_n true for odd n and false for even n.

4.10.7.2 B. Practice

Exercise. B1: Running the construction

Run the construction of Lemma 4.10.2 on $\Gamma = \{p \vee q,\ \neg p\}$ for the first four formulas of a list beginning $p,\ q,\ \neg q,\ p \wedge q$. Justify each "consistent" by a row and each "inconsistent" by a short derivation of $\bot$. Then say which values v_Δ gives p and q, whatever the rest of the list.

Solution. $\Delta_0 = \Gamma$.

p: Γ with p is inconsistent (p and $\neg p$, then $\neg$E), so Δ_1 adds $\neg p$, which is already there.

q: Δ_1 with q is consistent, since the row $p = \mathrm{F}$, $q = \mathrm{T}$ makes $p \vee q$, $\neg p$, q true (Section 4.7); Δ_2 adds q.

$\neg q$: Δ_2 with $\neg q$ is inconsistent (q and $\neg q$, then $\neg$E); Δ_3 adds $\neg\neg q$.

$p \wedge q$: Δ_3 with $p \wedge q$ is inconsistent ($\wedge$E gives p, which clashes with $\neg p$); Δ_4 adds $\neg(p \wedge q)$.

Whatever the rest of the list, $\neg p$ and q are in Δ, so $p \notin \Delta$ by Lemma 4.10.3 (b), and v_Δ makes p false and q true. Later stages only add formulas, so they cannot change this.

Marking scheme. (6 points)

- 4: one per stage, with its justification (a row checked, or a derivation of $\bot$).
- 2: the row of v_Δ on p, q, and why later stages cannot change it.

Exercise. B2: The clause for implication

Prove the $\to$ clause of Lemma 4.10.3 (c) in full, from parts (a) and (b) and the sequents of Lemma 4.8.1.

Solution. ($\Rightarrow$) Let $\varphi \to \psi \in \Delta$. If $\varphi \in \Delta$, the two-premise list $\varphi \to \psi,\ \varphi$ of members of Δ derives ψ by one $\to$E line, so $\Delta \vdash \psi$, and $\psi \in \Delta$ by (a). So either $\varphi \notin \Delta$ or $\psi \in \Delta$.

($\Leftarrow$) If $\varphi \notin \Delta$, then $\neg\varphi \in \Delta$ by (b), and ($\to$1), $\neg\varphi \vdash \varphi \to \psi$, shows $\Delta \vdash \varphi \to \psi$; so $\varphi \to \psi \in \Delta$ by (a). If $\psi \in \Delta$, then ($\to$2), $\psi \vdash \varphi \to \psi$, does the same.

Marking scheme. (6 points)

- 3: ($\Rightarrow$), with (a) cited.
- 3: ($\Leftarrow$), both cases, with (b) used in the first.

4.10.7.3 C. Going deeper

Exercise. C1: Coloring an infinite graph

A graph on the vertices $1, 2, 3, \ldots$ is given by saying which pairs of vertices are joined. A **3-coloring** gives each vertex one of three colors so that joined vertices get different colors. Suppose every finite set of vertices, with the joins among them, can be 3-colored. Prove that the whole graph can. Use atoms $c_{n,i}$, read "vertex n gets color i" for $i = 1, 2, 3$, where the $c_{n,i}$ are distinct atoms (for instance, $c_{n,i}$ is p_{3n+i}).

Solution. Let Γ consist of three kinds of formulas:

- $c_{n,1} \vee c_{n,2} \vee c_{n,3}$, for every vertex n;
- $\neg(c_{n,i} \wedge c_{n,j})$, for every vertex n and all $i \neq j$;
- $\neg(c_{m,i} \wedge c_{n,i})$, for every joined pair m, n and every i.

A valuation making every member of Γ true gives each vertex exactly one color (the first two kinds), and joined vertices different colors (the third kind): it is a 3-coloring, coloring n with the i for which $c_{n,i}$ is true.

Every finite list Γ_0 of members of Γ is satisfiable. Its members mention finitely many vertices; call that finite set F. By hypothesis, F with the joins among its vertices has a 3-coloring κ. Let v make $c_{n,i}$ true exactly when $n \in F$ and $\kappa(n) = i$. Each member of Γ_0 mentions only vertices of F, and v makes it true: the first kind because κ colors every vertex of F, the second because it gives each one color, the third because κ colors joined vertices of F differently.

By Theorem 4.10.6, Γ is satisfiable, and a satisfying valuation is a 3-coloring of the whole graph.

Marking scheme. (10 points)

- 3: the collection Γ, with all three kinds.
- 2: a valuation making Γ true gives a 3-coloring.
- 4: every finite part satisfiable, with the restriction to the vertices it mentions.
- 1: Theorem 4.10.6 cited.

Fatal errors (0 for the step): coloring the vertices one at a time and claiming the choices can always be continued, with no argument; a "finite part" argument that colors only the vertices $1, \dots, n$ without checking that every member of Γ_0 mentions only them.

Exercise. C2: Challenge: Compactness without derivations

Suppose the atoms are listed $a_1, a_2, \dots$, and every finite list of members of Γ is satisfiable. Choose truth values for $a_1, a_2, \dots$ one at a time, keeping this property:

(P_n) every finite list of members of Γ is made true by some valuation that agrees with the choices made for $a_1, \dots, a_n$.

(a) Show that if P_n holds, a value for a_{n+1} can be chosen so that P_{n+1} holds.

(b) Show that the valuation v given by all the choices makes every member of Γ true.

(c) Which earlier results does (b) use?

Solution (Hint 1). For (a), suppose neither value works, and take one finite list witnessing each failure.

Solution. (a) P_0 is the hypothesis. Suppose P_n holds and neither value of a_{n+1} works. Then there is a finite list F_1 that no valuation agreeing with the choices and making a_{n+1} true satisfies, and a finite list F_2 that no valuation agreeing with the choices and making a_{n+1} false satisfies. By P_n, the finite list F_1, F_2 is satisfied by some valuation u agreeing with the choices for $a_1, \dots, a_n$. If u makes a_{n+1} true, it satisfies F_1; if false, it satisfies F_2. Either way a contradiction. So some value works: choose T if T works, and F otherwise. This gives P_{n+1}, and the rule defines the values by recursion on n.

(b) By (a) and Axiom 4.2.1, P_n holds for every n. Let $\gamma \in \Gamma$, with its atoms among $a_1, \dots, a_N$. By P_N, the one-member list γ is made true by some valuation agreeing with v on $a_1, \dots, a_N$. By Lemma 4.4.1, v makes γ true.

(c) Lemma 4.4.1, Axiom 4.2.1 (P_n for every n, from (a)), and a definition by recursion on n (each choice depends on the earlier ones). No derivation and no completeness theorem is used: this proves Theorem 4.10.6 directly.

Marking scheme. (10 points)

- 5: (a), with the joint list F_1, F_2 and both cases of u.
- 4: (b), with N chosen from the atoms of γ, P_N obtained from (a) by Axiom 4.2.1, and Lemma 4.4.1 cited.
- 1: (c).

Fatal errors (0 for the step): in (a), choosing a value that works for each finite list separately, possibly different for different lists.

Chapter 5

Predicate Logic and the Quantifier Rules

Every lock in this building has a key that opens it. Some key opens every lock in this building. The two sentences use the same words, and the second says far more: it promises a master key. The statements of analysis are made of "for every" and "there exists", and to a truth table each is a single atom: the argument "every positive real has a square root; 2 is positive; so 2 has a square root" has the form $p, q \therefore r$, which a refuting row breaks. This chapter opens the atoms: predicates, terms and the two quantifiers, then their rules and those of equality. Each quantifier rule has a side condition about names, and the side conditions are exactly the rules an analyst breaks when a δ secretly depends on an x; continuity and uniform continuity differ only in the order of two quantifiers. *Why does "uniformly continuous implies continuous" have a proof a few lines long, while its converse is false?* Chapter 6 reads the answer off the rules.

The chapter is the first-order twin of Chapters 1 and 2: one rule per section, each with its card, its gallery of instances and its "Not allowed" block.

What you need. Chapters 1–4. From Chapters 1–3: the rules of every connective, derivations with boxes and the scope rule (Definition 1.4.2), derived rules (Definition 2.3.1), and the search method of Section 3.1, which each rule section here extends by one step. From Chapter 4: formulas as objects (Definition 4.1.1), structural induction and recursion (Theorem 4.3.1, Theorem 4.3.4), and refuting rows (Corollary 4.7.3). School arithmetic and the order of $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ are used in examples as facts.

Time. About 45 hours on the core track (all of A, one B and one C per section); about 65 hours for everything.

Minimum before *Sets, Relations, and Cardinality*. Chapters 1–6 and 8, core; Sections 7.1–7.2 before its Chapter 3.

How to read this chapter. Every section is on the first reading, with its folds closed. The folds hold proofs about the language (substitution, renaming, definitions), longer error analyses and edge cases. They are not optional: later proofs may use them.

5.0.1 Roadmap

- **Predicates and Quantifiers.** "For every" and "there exists" in a pictured world and over numbers; bounded quantifiers; the order of quantifiers.
- **Terms, Formulas, Free Variables and Substitution.** The exact language, free and

bound letters, substitution, and capture.
- **Using "For Every": Instantiation.** ∀E at the term the goal needs; counter-readings.
- **Proving "For Every": Let x Be Arbitrary.** ∀I and the flagged box: a fresh name, nothing assumed about it.
- **Proving "There Exists": Exhibit a Witness.** ∃I, with the witness chosen when a line shows it.
- **Using "There Exists": Name a Witness.** ∃E and its box; moving ∃ past ∀; the barber.
- **Equality.** =I and =E, chains, Leibniz's Law, and the table of the whole calculus.
- **Unique Existence, Counting and Descriptions.** "Exactly one", counting with $\neq$, "the x such that", graphs of functions, and why a definition proves nothing new.

5.0.2 Named moves

- **Read the dependency off the order.** A variable depends only on those quantified to its left. (Section 5.1)
- **Instantiate at the term you need** (∀E). Read it off the goal. (Section 5.3)
- **Break it with a counter-reading.** Premises true, the line false. (Section 5.3)
- **Let x be arbitrary** (∀I). A fresh name, about which nothing is assumed. (Section 5.4)
- **Postpone the witness** (∃I). Choose the term when a line shows it. (Section 5.5)
- **Name the witness early** (∃E). "Choose x_0 with ...", and never generalize over it. (Section 5.6)
- **Fix, obtain, choose.** ∀I; ∀E, ∃E on the hypotheses; ∃I. Most proofs in analysis. (Section 5.6)
- **Existence and uniqueness, separately.** Prove "at least one" and "at most one" apart. (Section 5.8)

The four quantifier rules are on one card at the end of Section 5.6, and the whole calculus on the table at the end of Section 5.7.

Where this goes. *Sets, Relations, and Cardinality*, Chapter 1 develops sets, membership, the empty set, pairs and functions, used here in their everyday sense (Sections 5.1, 5.2, 5.7, 5.8). *Analysis on the Real Line*, Chapter 1 proves from the axioms of the real numbers the school facts used in examples: that $\sqrt{2}$ exists (Sections 5.1, 5.8) and the field and order facts behind mixed chains (Section 5.7).

5.0.3 Summary

The language and its substitution are exact, capture included. Each quantifier rule has a side condition about names: ∀E at a term free for the variable, ∀I over a fresh flag, ∃I at a witness a line shows, ∃E with a fresh name kept in its box. With them and the two rules for equality, $\exists x\,\forall y$ gives $\forall y\,\exists x$ and a counter-reading breaks the converse, equals share every property (Proposition 5.7.2), no relation has the barber's property (Theorem 5.6.3), and a definition proves nothing new about the old words (Proposition 5.8.2).

What is open: a stalled search shows nothing, and the counter-readings are not yet certified (Chapter 7). The definitions of analysis, four quantifiers deep, are still to be translated, negated and derived (Chapter 6).

5.0.4 Terms

- **bounded quantifier**: $\forall x \in A$, $\exists x \in A$ (Definition 5.1.1).

- **counter-reading**: premises true, the line false (Section 5.3).
- **counting with equality**: at least, at most, exactly n (Definition 5.8.2).
- **definite description**: "the x such that φ" (Definition 5.8.3).
- **definitional extension**: a new symbol, nothing new proved (Proposition 5.8.2).
- **equality rules** =I, =E: $t = t$; replace equals (Definition 5.7.1).
- **existential elimination** $\exists$E: name a witness, freshly (Definition 5.6.1).
- **existential introduction** $\exists$I: from an instance (Definition 5.5.1).
- **flag, flagged box**: the fresh name of $\forall$I (Definition 5.4.1).
- **formula**, first-order: connectives and quantifiers over atoms (Definition 5.2.2).
- **free for**: substituting captures no variable (Definition 5.2.6).
- **free variable**: a letter no quantifier binds (Definition 5.2.3).
- **Leibniz's Law**: equals share every property (Proposition 5.7.2).
- **sentence**: no free variable (Definition 5.2.4).
- **substitution** $\varphi[t/x]$: t for free x (Definition 5.2.5).
- **term**: a name for an object (Definition 5.2.1).
- **unique existence** $\exists!$: at least and at most one (Definition 5.8.1).
- **universal elimination** $\forall$E: any instance (Definition 5.3.1).

5.1 Predicates and Quantifiers

Chapters 1–4 treated an atomic statement as a closed box, and the statements of analysis keep their whole content inside the box: "for every $\varepsilon > 0$ there is an N such that ..." lives in the words "for every" and "there is" and in the letters they govern. This section opens the box: it introduces predicates, variables, domains and the two quantifiers $\forall$ and $\exists$, and teaches how to **read** a quantified sentence, first in a small picture and then over the numbers; the rules come one per section, in Sections 5.3 to 5.6.

5.1.1 An argument the truth table rejects

Here is an argument that nobody would doubt.

> Every positive real number has a square root. The number 2 is positive. Therefore 2 has a square root.

It is correct in the plainest sense: if the two premises are true, the conclusion must be. Now try to check it with the tools of Chapters 1–4. The three sentences contain no connective at all, so each is atomic. Give them atoms:

- p: "every positive real number has a square root";
- q: "2 is positive";
- r: "2 has a square root".

The argument has the form $p, q \therefore r$. Three different atoms, and nothing connects them. The valuation that makes p and q true and r false makes both premises true and the conclusion false. So by Corollary 4.7.3,

$$p,\ q \nvdash r.$$

No derivation, with any rule of Chapters 1 and 2, leads from p and q to r.

Second reading · *Why the truth table is not wrong*

There is no contradiction here, and it is worth seeing exactly why. Corollary 4.7.3 is about the **form** $p, q \therefore r$: it says that no derivation leads from two arbitrary statements to a third, unrelated one. That is true. Our argument is correct for a reason that the form cannot see. The first premise speaks about **every** positive real number. The second says that 2 is **one** of the things the first premise speaks about. The conclusion says of 2 what the first premise said of all of them. The three statements share parts: a property ("is positive"), another property ("has a square root"), an object (2), and the word "every". Chopped into atoms, the shared parts disappear.

This is the warning of Section 4.4 come true: a truth table sees only the connectives, and everything inside an atom is invisible to it. Section 4.4 closed with the remark that this chapter looks inside atoms. We now do. The first job is a vocabulary for the parts inside: the properties, the objects, and the words "every" and "some". Section 5.3 then gives the rule that makes the argument above a derivation, four lines long.

5.1.2 Predicates: statements with holes

Chapter 1 met the sentence "$x > 2$" and called it an **open sentence** (item (e) of Example (Statements and nonstatements)). It is not a statement, because its truth depends on what x is: it is true when x is 3 and false when x is 0. Put a number in for x and it becomes a statement.

The argument of the last subsection is built from two open sentences of this kind: "x is positive" and "x has a square root". Each has one hole, marked by x. The premise "2 is positive" is the first open sentence with 2 put into the hole; the conclusion is the second, with 2 put into its hole. The first premise does something else with the holes, which we come to in the next subsection. Open sentences are the missing parts, and they deserve a name.

A predicate is a statement with holes in it; filling every hole with a name of an object gives a statement.

A **predicate** is an open sentence with one or more marked holes. The number of holes is its **arity**: a predicate with one hole is **one-place** (or unary), with two holes **two-place** (binary), and so on. We write a predicate as a capital letter followed by its holes in parentheses: $P(x)$ for "x is positive", $S(x)$ for "x has a square root", $D(x, y)$ for "x divides y". The letters that mark the holes are **variables**: x, y, z, u, and in analysis also $\varepsilon, \delta, N, n$. A **constant** is a name of one particular object: 0, 1, 2, or a letter a, b, c that has been given a fixed meaning.

In words: a predicate says something **about** objects without saying which objects. The variables are placeholders, not objects. "x is positive" is about whatever will be put in for x, and until something is, it is neither true nor false.

Many predicates of mathematics are written with a symbol between their two holes rather than a letter in front: $x < y$, $x \leq y$, $x = y$, $x \mid y$. These are two-place predicates like any other; only the way of writing them differs. Membership $x \in A$, for a fixed set A, is a one-place predicate of x: "x is an element of A". Here a set is understood informally, as a collection of objects (usually numbers), and nothing in this section uses more about a set A than whether a given object is an element of it. For sets A and B, $A \subseteq B$ means that every element of A is an element of B; so $A \subseteq \mathbb{R}$ says that A is a set of real numbers.

Objects can also be built from other objects: $x+1$, x^2, $|x|$, $2 \cdot y$. The symbols that build them ($+$, squaring, $|\cdot|$, $\cdot$) are **function symbols**, and an expression such as x^2+1 names an object as soon as x does. A function symbol may also be a letter, $f(x)$. Here a function symbol is simply a way of naming one object in terms of others. Section 5.2 makes all of this exact, under the name **term**. For now, a term is any expression that names an object: a variable, a constant, or a function symbol applied to terms.

Filling holes is the first way to turn a predicate into a statement.

- $P(2)$: "2 is positive". A statement, and true.
- $P(-3)$: "-3 is positive". A statement, and false.
- $D(3, 12)$: "3 divides 12". True.
- $D(3, y)$: "3 divides y". Still open: one hole is filled and one is not. It is a one-place predicate of y.
- $P(x^2+1)$: "x^2+1 is positive". Still open, in x. (It happens to be true whatever real number x is, but that is a fact about the numbers, and the sentence still has a hole.)

The last two items show that filling holes is not all or nothing. A predicate with some holes filled is a predicate with fewer holes. Only when **every** hole is filled, or closed in the way described next, do we have a statement.

Some letters change their job in this chapter:

Letter	In Chapters 1–4	In this chapter	How to tell
P, Q, R	a whole statement (Chapters 1–2); an atom or any formula (Chapter 4)	a predicate symbol	followed by its holes: $P(x)$, $R(x, y)$; a bare P is a statement letter from earlier chapters
p, q, r	atoms of a truth table	atoms still, only where a truth table reads a formula (the argument above; review exercises)	never a variable or a symbol of a first-order formula in this chapter
x, y, n, ε	one fixed number in an example ("let x be a real number")	variables: holes of predicates, closed by quantifiers	a quantifier in front of them, or a hole still open

A property $P(n)$ of natural numbers, as in the principle of induction (Axiom 4.2.1), was already a one-place predicate in this sense.

5.1.3 Quantifiers and the domain

There is a second way to turn "x is positive" into a statement, and the first premise of our argument uses it. The premise does not put a particular number into the hole of "x has a square root". It says that the open sentence "if x is positive, then x has a square root" is true **whatever** real number is put in. The word "every" closes the hole without filling it.

Two questions have to be answered before "whatever is put in" makes sense. What may be put in? And how many of those things must make the sentence true?

The first question is answered by a **domain of discourse**, or **domain**: the collection of objects

that the variables range over. For the argument above, the domain is the real numbers, $\mathbb{R}$. In number theory it is often $\mathbb{Z}$ or $\mathbb{N}$ (in this book $\mathbb{N} = \{0, 1, 2, \dots\}$, with 0). A domain always has at least one element; an empty collection is handled by bounded quantifiers, as below.

The second question is answered by the **quantifier**.

"For every" promises something of each object in the domain; "there exists" promises it of at least one.

Let φ be an open sentence in which the variable x marks a hole (it may have other holes too). There are two quantifiers.

- The **universal quantifier**: $\forall x\, \varphi$, read "**for every** x, φ", "for all x, φ", or "for each x, φ". It is **true** when **every** element of the domain, put in for x, makes φ true.
- The **existential quantifier**: $\exists x\, \varphi$, read "**there exists** an x such that φ", "there is an x with φ", or "for **some** x, φ". It is **true** when **at least one** element of the domain, put in for x, makes φ true.

In words: $\forall x$ and $\exists x$ each close the hole marked x, so they remove x from the list of holes. $\forall x\, \varphi$ makes one claim about the whole domain at once: no element is an exception. $\exists x\, \varphi$ claims that there is at least one element for which φ holds, and it does not say which, or how many. "At least one" includes the case of exactly one and the case of all of them. The word "some" in mathematics never means "some but not all".

With these, the first premise of our argument, over the domain $\mathbb{R}$, is

$$\forall x\, (P(x) \to S(x)),$$

"for every x, if x is positive, then x has a square root". The second premise is $P(2)$ and the conclusion is $S(2)$. Now the shared parts are visible: P occurs in both premises, S in the first premise and the conclusion, and 2 in the second premise and the conclusion. The rule of Section 5.3 uses the first premise with x replaced by 2, giving $P(2) \to S(2)$, and $\to$E finishes.

Notice the $\to$ inside the quantifier. The premise does not say that every real number is positive and has a square root; it says that every real number which is positive has one. The subsection on bounded quantifiers below explains why a "for every" restricted to part of the domain always comes with an implication.

Remark. "True" here means what it means in ordinary mathematics: look at the domain and ask whether every element, or some element, makes the sentence true. A claim that a sentence can be **derived** is always proved by a derivation, never by this reading.

5.1.3.1 Quantifiers in a small world

Before the numbers, a picture. A **world** here is a few dots, the objects of the domain, with some dots labeled P or Q (the dots the one-place predicates hold of) and arrows between dots (the pairs a two-place predicate R holds of: an arrow from u to v means $R(u, v)$).

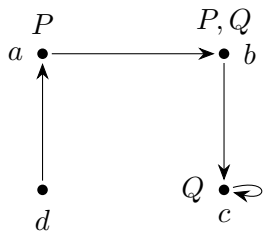


Figure W. World W: domain $\{a, b, c, d\}$; P holds of a, b; Q of b, c; R is the four arrows.

Six readings, each checked dot by dot:

- $\exists x\, P(x)$: true; the witness a.
- $\forall x\, P(x)$: false; the counterexample c (or d).
- $\forall x\, (P(x) \vee Q(x))$: false at d, which has neither label.
- $\exists x\, (P(x) \wedge Q(x))$: true at b.
- $\forall x\, \exists y\, R(x, y)$: true; every dot has an arrow leaving it ($a \to b$, $b \to c$, $c \to c$, $d \to a$).
- $\exists y\, \forall x\, R(x, y)$: false; no dot receives an arrow from all four (c receives from b and c only).

In a finite world every quantifier can be checked dot by dot. Over $\mathbb{N}$ or $\mathbb{R}$ it cannot, and the reading needs an argument, as the next subsection shows.

Quick check. In world W, is $\forall x\, (Q(x) \to R(x, c))$ true? Is $\exists x\, \forall y\, \neg R(y, x)$ true?

Solution. The first is true: Q holds of b and c, and both have an arrow to c; at a and d the hypothesis $Q(x)$ is false, so the implication holds. The second is true, with the witness d: no arrow ends at d.

A long "and", a long "or". When the dots of a world can be listed, a quantifier is a connective repeated. In world W, $\forall x\, P(x)$ says the same as $P(a) \wedge P(b) \wedge P(c) \wedge P(d)$, and $\exists x\, P(x)$ the same as $P(a) \vee P(b) \vee P(c) \vee P(d)$. A counterexample to the first is a false conjunct (c); a witness for the second is a true disjunct (a). De Morgan's law (Proposition 2.5.5), used three times, then says that "not every dot has P" is "some dot lacks P": $\neg(P(a) \wedge \cdots \wedge P(d))$ and $\neg P(a) \vee \cdots \vee \neg P(d)$. Over $\mathbb{N}$ the conjunction would never end, and no formula can write it down; that is why the quantifiers get rules of their own.

Quick check. In world W, write $\forall x\, (P(x) \to Q(x))$ as a conjunction of four statements, one per dot. Which conjunct is false?

Solution. $(P(a) \to Q(a)) \wedge (P(b) \to Q(b)) \wedge (P(c) \to Q(c)) \wedge (P(d) \to Q(d))$. The first conjunct is false, since a has P and not Q. The other three are true: b has both labels, and c and d lack P.

5.1.3.2 The same sentence over different domains

A quantified sentence is a claim about its domain, so changing the domain can change its truth value. Two examples show how.

Take $\forall x\, (x \geq 0)$, "every x is at least 0".

- Over $\mathbb{N}$ it is **true**: every natural number is at least 0.
- Over $\mathbb{Z}$ it is **false**: the integer -1 is in the domain, and $-1 \geq 0$ is false.

One element of the domain that makes φ false is enough to make $\forall x\, \varphi$ false. Such an element is a **counterexample**. To show that the sentence is true over $\mathbb{N}$, on the other hand, we had to say something about every natural number at once.

Take $\exists x\,(x^2 = 2)$, "some x has square 2".

- Over $\mathbb{R}$ it is **true**: the real number $\sqrt{2}$ has square 2. (Its existence is a school fact here.)
- Over $\mathbb{Q}$ it is **false**: that no rational number has square 2 is exactly Theorem 2.2.1.

One element that makes φ true is enough to make $\exists x\,\varphi$ true. Such an element is a **witness**. To show that the sentence is false over $\mathbb{Q}$, on the other hand, a theorem about every rational number was needed, and Chapter 2 proved it by contradiction.

The two quantifiers are mirror images. For $\forall$, a single counterexample refutes and a general argument confirms. For $\exists$, a single witness confirms and a general argument refutes.

Quick check. Over which of the domains $\mathbb{Z}_{>0} = \{1, 2, 3, \dots\}$, $\mathbb{Z}$, $\mathbb{R}$ is the sentence $\exists x\,(x + x = x)$ true? Give a witness where it is true, and a reason where it is false.

Solution. Only $x = 0$ satisfies $x + x = x$, since subtracting x from both sides gives $x = 0$. So the sentence is true over $\mathbb{Z}$ and over $\mathbb{R}$, with the witness 0. It is false over $\mathbb{Z}_{>0}$, which does not contain 0: every $n \in \mathbb{Z}_{>0}$ has $n + n > n$, since $n > 0$. (Over $\mathbb{N} = \{0, 1, 2, \dots\}$ it would be true, again with the witness 0.)

5.1.3.3 Reading a quantified sentence

Quantified sentences are read from the outside in: the first quantifier governs everything after it, and inside it may sit another quantifier, a connective, or a predicate. Reading aloud slowly, one symbol at a time, is the most reliable way to understand one. Here are three, simplest first.

Example. Reading three sentences

Read each sentence aloud, say what it claims, and decide whether it is true over the domain given.

(a) $\forall x\,(x^2 \geq 0)$, over $\mathbb{R}$.

(b) $\exists x\,\forall y\,(x \leq y)$, over $\mathbb{N}$.

(c) $\neg\exists m\,(\text{Prime}(m) \wedge \forall n\,(\text{Prime}(n) \to n \leq m))$, over $\mathbb{N}$, where $\text{Prime}(m)$ is the one-place predicate "m is prime".

Solution. (a) "For every x, x^2 is at least 0." The claim is about every real number, and it is true: the square of a real number is never negative (a school fact about signs: $x \cdot x$ is a product of two numbers of the same sign, or 0).

(b) "There is an x such that for every y, $x \leq y$." Unpack from the outside. The outer quantifier $\exists x$ asks for one natural number x. What must be true of it is the inside, $\forall y\,(x \leq y)$: every natural number y is at least x. So the sentence says that $\mathbb{N}$ has a **least element**. It is true, with the witness $x = 0$: every natural number is at least 0.

(c) The outer symbol is $\neg$, so the sentence denies what follows. What follows is $\exists m\,(\dots)$: there is an m that is prime and such that every prime n is at most m. Such an m would be a **largest prime**. So the sentence says "there is no largest prime". It is true: this is the ancient theorem that there are infinitely many primes. Nothing in this book depends on that theorem; here it is only something to read.

Notice the shape inside (c): $\forall n\,(\text{Prime}(n) \to n \leq m)$, "every prime is at most m", has an implication under the $\forall$. It speaks about primes only, and the implication is how a "for every" is restricted to them. The next subsection makes this a rule of writing.

The sentence in (b) has two quantifiers, and they come in a definite order: first $\exists x$, then $\forall y$. The order matters, and the subsection "The order of quantifiers" below is about it. First, restricted quantifiers.

5.1.4 Bounded quantifiers

Mathematics rarely says "for every real number". It says "for every $\varepsilon > 0$", "for every $x \in A$", "there is an $n \geq N$". The quantifier ranges over a **part** of the domain, described by a condition. There is no new quantifier for this; the condition is written into the sentence with a connective. Which connective turns out to matter.

A "for every" restricted to a condition is an implication; a "there exists" restricted to a condition is a conjunction.

Definition 5.1.1. Bounded Quantifier

Let A be a set, x a variable, and φ an open sentence. The **bounded quantifiers** are the abbreviations

$$\forall x \in A\ \varphi \quad \text{stands for} \quad \forall x\,(x \in A \to \varphi),$$
$$\exists x \in A\ \varphi \quad \text{stands for} \quad \exists x\,(x \in A \wedge \varphi).$$

In the same way, for a term a in which x does not occur, $\forall x > a\ \varphi$ stands for $\forall x\,(x > a \to \varphi)$ and $\exists x > a\ \varphi$ for $\exists x\,(x > a \wedge \varphi)$, and likewise with $\geq$, $<$, $\leq$ in place of $>$.

In words: "for every x in A, φ" means "for every x whatsoever, **if** x is in A, **then** φ". "There is an x in A with φ" means "there is an x that is in A **and** satisfies φ". The condition $x \in A$, or $x > a$, is the **bound**. A bounded quantifier is not a new symbol of the language: it is shorthand, and every statement about it is a statement about the formula it stands for.

↻ **Second reading** · *The induction display, read again*

The display after Axiom 4.2.1, in Section 4.2, used $\forall n \in \mathbb{N}$ as an abbreviation of the words "for every natural number n". It was already this bounded form: $\forall n \in \mathbb{N}\ \big(P(n) \to P(n+1)\big)$ stands for $\forall n\,(n \in \mathbb{N} \to (P(n) \to P(n+1)))$.

Why $\to$ under $\forall$, and $\wedge$ under $\exists$. Try the other choices and see what they say. Take the domain $\mathbb{R}$, the set $A = \{1, 2\}$, and the open sentence $x > 5$.

- "Every element of A is greater than 5" is false: 1 is not. The correct form $\forall x\,(x \in A \to x > 5)$ is false too, with the counterexample $x = 1$, which is in A and not greater than 5. The wrong form $\forall x\,(x \in A \wedge x > 5)$ says that **every real number** is in A and greater than 5. It is false as well, but for a reason that has nothing to do with A's elements: $x = 0$ is not in A. Change A to $\{6, 7\}$ and the correct form becomes true while the wrong one stays false.

- "Some element of A is greater than 5" is false for $A = \{1, 2\}$. The correct form $\exists x\,(x \in A \wedge x > 5)$ is false too: no real number is both in A and greater than 5. The wrong form $\exists x\,(x \in A \to x > 5)$ is **true**, with the witness $x = 0$: since $0 \notin A$, the implication "if $0 \in A$ then $0 > 5$" has a false hypothesis and is true (vacuous truth, the phenomenon behind Proposition 2.1.2).

So a $\forall$ with $\wedge$ inside claims too much: it says the whole domain lies in A. An $\exists$ with $\to$ inside claims too little: anything outside A is a witness. Neither says what was meant.

Warning. An existential quantifier with an implication directly inside, $\exists x\,(\alpha \to \beta)$, is almost always a mistranslation. Over $\mathbb{R}$, "some element of $\{1, 2\}$ exceeds 5" is false, but $\exists x\,(x \in \{1, 2\} \to x > 5)$ is true, because $x = 0$ makes the hypothesis false. When you write $\exists$ and then $\to$, check whether you meant $\wedge$.

Mixing kinds of number. A sentence of analysis often mixes real numbers and natural numbers; the convention is to take the larger kind as the domain and let the smaller kind enter as a bound, as in $\forall \varepsilon > 0\ \exists N \in \mathbb{N}$, with a sequence (a_n) read as a function symbol a, $a_n = a(n)$.

The empty set. Bounded quantifiers make sense even when the bound is never satisfied. Let $\emptyset$ be the **empty set**, the set with no elements, so that $x \in \emptyset$ is false for every x.

- $\forall x \in \emptyset\ \varphi$ stands for $\forall x\,(x \in \emptyset \to \varphi)$. For every x, the hypothesis $x \in \emptyset$ is false, so the implication is true. The sentence is **true**, whatever φ says. For example, "every element of $\emptyset$ is greater than 5" and "every element of $\emptyset$ is less than 5" are both true.
- $\exists x \in \emptyset\ \varphi$ stands for $\exists x\,(x \in \emptyset \wedge \varphi)$. No x makes the first conjunct true, so the sentence is **false**, whatever φ says.

Section 2.1 met this phenomenon for implications: a statement whose hypothesis never holds is true, and says nothing. A universal statement about the elements of an empty set is true for the same reason, because there is nothing to test. This is the reading; Section 5.4 derives the universal half with the rules.

Warning. **Existential import.** "Every element of A has property φ" does **not** imply "some element of A has property φ". A "for every" over a bound has no *existential import*: it does not say that anything satisfies the bound. It fails when A is empty: "every element of $\emptyset$ is greater than 5" is true, and "some element of $\emptyset$ is greater than 5" is false. The step from "every" to "some" needs an element, and the bound can remove all of them.

5.1.5 The order of quantifiers

A sentence with two quantifiers of different kinds can change its meaning completely when they are swapped. This is the single most important fact about reading quantifiers, and the definitions of analysis depend on it.

Example. Swapping two quantifiers

Over the domain $\mathbb{R}$, compare

$$\text{(i)}\ \ \forall x\,\exists y\,(x + y = 0) \qquad \text{and} \qquad \text{(ii)}\ \ \exists y\,\forall x\,(x + y = 0).$$

Read each aloud, and decide which is true.

Solution. *(i)* "For every x there is a y with $x + y = 0$": every real number has an additive inverse. It is **true**. Given any real x, the number $y = -x$ satisfies $x + y = x + (-x) = 0$.

(ii) "There is a y such that for every x, $x + y = 0$": one single number y is an inverse of every real number at once. It is **false**. Whatever real number y is proposed, take $x = 1 - y$; then $x + y = 1 \neq 0$. So no y works.

The two sentences have the same letters, the same quantifiers and the same equation. Only the order differs. In (i), the y may be chosen **after** x is known, so it may depend on x: the witness $-x$ is a different number for each x. In (ii), y must be chosen **first**, once and for all, and then it must work for every x that comes after.

Second reading · *What makes the pair work*

What makes the pair work as an example? Two properties, and any other pair with the same shape would do as well. First, for each x separately there is a suitable y. Second, no one y is suitable for all x together, because each y serves only one x (here, only $x = -y$).

Quantifiers as a game. Read a sentence as a game between two players, the challenger, who tries to make it false, and the defender, who tries to make it true.

- Each $\forall$ is a move of the challenger, who picks a value; each $\exists$ is a move of the defender.
- The moves are made in the order the quantifiers are written, left to right, and each player sees every move made so far.
- When every variable has a value, the inside is checked. The sentence is true exactly when the defender can win whatever the challenger does.

Here are the two sentences of Example (Swapping two quantifiers), played.

	(i) $\forall x\, \exists y\, (x + y = 0)$	(ii) $\exists y\, \forall x\, (x + y = 0)$
first move	challenger picks x, say $x = 5$	defender picks y without knowing x, say $y = 3$
second move	defender answers $y = -x = -5$	challenger answers $x = 1 - y = -2$
check	$5 + (-5) = 0$: defender wins	$-2 + 3 = 1 \neq 0$: challenger wins
who always wins	the defender, with $y = -x$	the challenger, with $x = 1 - y$

Remark. The game reads sentences whose quantifiers all stand at the front, unnegated, before an inside with no quantifiers. A bounded quantifier such as $\forall \varepsilon > 0$ counts, its bound checked with the inside. Under a $\neg$, or in the hypothesis of an implication, a quantifier changes sides.

Read the dependency off the order: a variable may depend on the variables quantified to its left,

and only on them. In (i), y is to the right of x, so the choice of y may use x. In (ii), y is to the left, so it may not. Sections 5.4 to 5.6 will show this same dependence as a fact about the boxes of a derivation.

Quick check. Over the domain $\mathbb{N}$, swap the two quantifiers in "for every n there is a prime larger than n", that is, in $\forall n\, \exists m\, (\text{Prime}(m) \wedge m > n)$. Read the result aloud. Is it true?

Solution. The swapped sentence is $\exists m\, \forall n\, (\text{Prime}(m) \wedge m > n)$: "there is a prime that is larger than every natural number". It is **false**. Whatever m is proposed, take $n = m$: then $m > n$ is false, so the inside fails. (The original sentence is true; it is another form of the theorem that there are infinitely many primes, which we do not use.) As in Example (Swapping two quantifiers), the original lets the prime depend on n, and the swapped sentence demands one prime for all n at once.

↻ **Second reading** · *When swapping is harmless*

When swapping is harmless. Two quantifiers of the **same** kind may be swapped. $\forall x\, \forall y\, \varphi$ and $\forall y\, \forall x\, \varphi$ both say that φ holds for every choice of the pair; in the game, the challenger makes both moves, and the order of the challenger's own moves gives nothing away. Likewise $\exists x\, \exists y\, \varphi$ and $\exists y\, \exists x\, \varphi$ both say that some pair works. Section 5.4 derives the first swap with the rules. For quantifiers of **different** kinds, one direction is always safe: if one y works for every x, then for every x some y works (the same one). Section 5.6 derives this, $\exists y\, \forall x\, \varphi$ to $\forall x\, \exists y\, \varphi$, and Example (Swapping two quantifiers) shows that the other direction fails.

5.1.6 Words that hide quantifiers

English marks quantifiers loosely, and mathematical English has habits that hide them altogether. Two traps are worth naming now, before they cause trouble.

Warning. The word **"any"** can mean "every" or "some", and only the sentence around it decides. "Any x in A is positive" means every element of A is positive: $\forall x \in A\ (x > 0)$. "If any x in A is positive, then A is non-empty" means if **some** element is positive: $\exists x \in A\ (x > 0) \to \ldots$. When you meet "any", rewrite the sentence with "every" and with "some", and keep the one that says what was meant.

Hidden universal quantifiers. A theorem stated as "if n is even, then n^2 is even" has no quantifier word in it, and yet it is not about one number. It means: for every integer n, if n is even, then n^2 is even. A letter that is not introduced anywhere, in the statement of a theorem, is read with a "for every" in front. When translating, put the hidden quantifier back first.

5.1.7 Keys and locks

A domain with no numbers in it makes the shapes easy to see: the objects of a building, with $\text{Key}(x)$: "x is a key", $\text{Lock}(x)$: "x is a lock", and $\text{Opens}(x, y)$: "x opens y". The table reads in either direction; c is a constant, "the front lock".

English	Formula	Note
every key opens lock c	$\forall x\,(\text{Key}(x) \to \text{Opens}(x, c))$	$\to$ under $\forall$
some key opens lock c	$\exists x\,(\text{Key}(x) \wedge \text{Opens}(x, c))$	$\wedge$ under $\exists$
no key opens lock c	$\forall x\,(\text{Key}(x) \to \neg\text{Opens}(x, c))$, or $\neg\exists x\,(\text{Key}(x) \wedge \text{Opens}(x, c))$	two equally good forms
only keys open lock c	$\forall x\,(\text{Opens}(x, c) \to \text{Key}(x))$	"only" turns the arrow round
every lock has a key that opens it	$\forall y\,(\text{Lock}(y) \to \exists x\,(\text{Key}(x) \wedge \text{Opens}(x, y)))$	every … some
some key opens every lock	$\exists x\,(\text{Key}(x) \wedge \forall y\,(\text{Lock}(y) \to \text{Opens}(x, y)))$	some … every

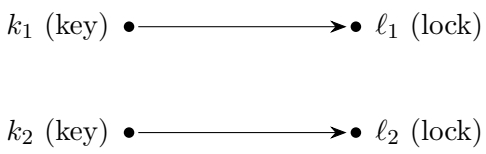


Figure K. A building where every lock has a key that opens it, and no key opens every lock.

In this building the fifth sentence of the table is true (ℓ_1 has k_1, ℓ_2 has k_2) and the sixth is false (k_1 fails at ℓ_2, k_2 at ℓ_1). The sixth implies the fifth, since a master key opens each lock, so the order "some … every" says more, exactly as in Example (Swapping two quantifiers).

Quick check. Translate "no key opens every lock". Is it true in the building of Figure K?

Solution. $\neg\exists x\,(\text{Key}(x) \wedge \forall y\,(\text{Lock}(y) \to \text{Opens}(x, y)))$, or $\forall x\,(\text{Key}(x) \to \neg\forall y\,(\text{Lock}(y) \to \text{Opens}(x, y)))$. It is true in the building of Figure K, since k_1 does not open ℓ_2 and k_2 does not open ℓ_1.

Second reading · *Two slips with an implication inside "there is"*

"Some key opens lock c" is **not** $\exists x\,(\text{Key}(x) \to \text{Opens}(x, c))$: that formula is true in every building, since the lock c is not a key, so c itself makes the hypothesis $\text{Key}(c)$ false. It is the warning after Definition 5.1.1 again.

An implication inside $\exists$. This is the warning after Definition 5.1.1 once more, and it comes up in analysis: "there is an N such that, if $n \geq N$, then …" is not the shape of "for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that for every $n \geq N$, …", where the implication $n \geq N \to \ldots$ sits under $\forall n \in \mathbb{N}$. The implication belongs under a $\forall$.

5.1.8 What comes next

We can read quantified sentences, and translate into them. We cannot yet reason with them. Section 5.2 makes the language exact, because every rule will put a term in for a variable; Sections 5.3–5.6 give the four rules.

Remark (Looking ahead). Chapter 7 turns "true in a domain" into a definition, and proves that a reading in which the premises are true and the conclusion false rules out every derivation.

5.1.9 Exercises

5.1.9.1 A. Check your understanding

Exercise. A1

(a) Read each sentence aloud, and say over which of the domains $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ it is true: (i) $\forall x\,(x{+}1 > x)$; (ii) $\exists x\,(x{+}1 = 0)$; (iii) $\forall x\,\exists y\,(y < x)$; (iv) $\exists x\,(2{\cdot}x = 1)$; (v) $\forall x\,(x{\cdot}x \geq x)$.

(b) With domain $\mathbb{R}$, which of the following translates "some element of A is negative"? (i) $\exists x\,(x \in A \to x < 0)$; (ii) $\exists x\,(x \in A \wedge x < 0)$; (iii) $\forall x\,(x \in A \wedge x < 0)$. For each of the two others, say what it claims instead.

(c) What is the arity of the predicate "$x + y = z$", and how many holes does "$3 + y = z$" have left?

(d) True or false, with a reason: $\forall x \in \emptyset\ (x \neq x)$ is true.

Solution. (a) (i) "For every x, $x + 1$ is greater than x." True over all three domains: adding 1 always gives a larger number.

(ii) "There is an x with $x + 1 = 0$." The only solution is $x = -1$. False over $\mathbb{N}$, which does not contain -1; true over $\mathbb{Z}$ and $\mathbb{R}$, with the witness -1.

(iii) "For every x there is a y smaller than x." False over $\mathbb{N}$: for $x = 0$ there is no natural number $y < 0$. True over $\mathbb{Z}$ and $\mathbb{R}$: given x, take $y = x - 1$.

(iv) "There is an x with $2x = 1$." The only solution is $x = \frac{1}{2}$. False over $\mathbb{N}$ and $\mathbb{Z}$, true over $\mathbb{R}$ with the witness $\frac{1}{2}$.

(v) "For every x, $x \cdot x$ is at least x." True over $\mathbb{N}$: if $x = 0$, then $x \cdot x = 0 = x$; if $x \geq 1$, then $x \cdot x \geq 1 \cdot x = x$. True over $\mathbb{Z}$: an integer is either at least 1, as before, or at most 0, and then $x \cdot x \geq 0 \geq x$. False over $\mathbb{R}$: the counterexample $x = \frac{1}{2}$ has $x \cdot x = \frac{1}{4} < \frac{1}{2}$.

(b) (ii) is the translation (Definition 5.1.1: $\wedge$ under $\exists$). (i) is true as soon as some real number lies outside A, since that number makes the implication vacuously true; so it can be true when no element of A is negative (for $A = \{1\}$, say). (iii) claims that every real number is in A and negative, which is false whatever A is (take $x = 0$).

(c) It has three holes, x, y, z, so it is three-place. Filling x with 3 leaves two holes, y and z.

(d) True. It stands for $\forall x\,(x \in \emptyset \to x \neq x)$. For every x the hypothesis $x \in \emptyset$ is false, so the implication is true, even though its conclusion $x \neq x$ is false for every x.

Exercise. A2: Drill

Answer each part in a line or two. Parts (a)–(d) are about world W of Figure W, parts (e)–(i) about the building with Key, Lock and Opens, and parts (j)–(l) about numbers.

(a) In W, is $\forall x\,(P(x) \to \exists y\,R(x,y))$ true?
(b) In W, is $\exists x\,(Q(x) \wedge \neg P(x))$ true?
(c) In W, is $\forall x\,\exists y\,R(y,x)$ true?
(d) In W, write $\exists x\,R(x,x)$ as a disjunction of four statements, one per dot, and decide it.
(e) Translate "every key opens some lock".
(f) Translate "no lock is a key".
(g) Translate "only keys open locks".
(h) Translate "some lock is opened by every key".
(i) Is (h) true in the building of Figure K?
(j) Over $\mathbb{Z}$, translate "every integer has a larger one", and say whether it is true.
(k) Swap the quantifiers of (j) and read the result.
(l) Over $\mathbb{R}$, translate with bounded quantifiers "some positive number is smaller than every positive number", and say whether it is true.

Solution.

(a) Yes: P holds of a and b, and $a \to b$, $b \to c$.
(b) Yes, at c.
(c) No: no arrow ends at d.
(d) $R(a,a) \vee R(b,b) \vee R(c,c) \vee R(d,d)$; true, by the third disjunct (the loop at c).
(e) $\forall x\,(\mathrm{Key}(x) \to \exists y\,(\mathrm{Lock}(y) \wedge \mathrm{Opens}(x,y)))$.
(f) $\forall x\,(\mathrm{Lock}(x) \to \neg\mathrm{Key}(x))$, or $\neg\exists x\,(\mathrm{Lock}(x) \wedge \mathrm{Key}(x))$.
(g) $\forall x\,\forall y\,((\mathrm{Lock}(y) \wedge \mathrm{Opens}(x,y)) \to \mathrm{Key}(x))$.
(h) $\exists y\,(\mathrm{Lock}(y) \wedge \forall x\,(\mathrm{Key}(x) \to \mathrm{Opens}(x,y)))$.
(i) No: ℓ_1 is not opened by k_2, and ℓ_2 not by k_1.
(j) $\forall x\,\exists y\,(y > x)$; true, with $y = x + 1$.
(k) $\exists y\,\forall x\,(y > x)$, "some integer is larger than every integer"; false: $x = y$ defeats every y.
(l) $\exists x > 0\;\forall y > 0\;(x < y)$; false: for any proposed $x > 0$, $y = x$ gives $x < x$.

5.1.9.2 B. Practice

Exercise. B1: Ordered quantifiers

Determine which of the following sentences are true over the domain given. Justify your answer: give the witness a true $\exists$ asks for (as a function of the earlier variables, if there are any), or the counterexample that refutes a $\forall$.

(a) $\forall x\,\exists y\,(x < y)$, over $\mathbb{N}$.
(b) $\forall x\,\exists y\,(x + y = 1)$, over $\mathbb{N}$.

(c) $\exists x\, \forall y\, (x \cdot y = y)$, over $\mathbb{R}$.

(d) $\forall x\, \exists y\, (y \cdot y = x)$, over $\mathbb{R}$.

Solution. *(a)* True. Given any natural number x, the natural number $y = x + 1$ satisfies $x < y$. The witness depends on x, which the order $\forall x\, \exists y$ allows.

(b) False. The challenger chooses $x = 2$. Then $x + y = 1$ forces $y = -1$, which is not a natural number; every natural y gives $2 + y \geq 2$. So for $x = 2$ no y works. (For $x = 0$ and $x = 1$ the witnesses $y = 1$ and $y = 0$ do exist, so the challenger must choose with care.) (Over $\mathbb{Z}$ the sentence would be true, with $y = 1 - x$.)

(c) True, with the witness $x = 1$: for every real y, $1 \cdot y = y$. The witness is chosen first and does not depend on y, as the order $\exists x\, \forall y$ demands.

(d) False, with the counterexample $x = -1$: for every real y, $y \cdot y \geq 0 > -1$, so no y has $y \cdot y = -1$.

Exercise. B2: Order matters

Give an example of a two-place predicate $R(x, y)$ on the domain $\mathbb{N}$ such that $\forall x\, \exists y\, R(x, y)$ is true and $\exists y\, \forall x\, R(x, y)$ is false. Check both claims.

Solution. Take $R(x, y)$: "$y > x$".

$\forall x\, \exists y\, (y > x)$ *is true.* Given any natural number x, the natural number $y = x + 1$ satisfies $y > x$.

$\exists y\, \forall x\, (y > x)$ *is false.* Whatever natural number y is proposed, take $x = y$: then $y > x$ is false. So no y is greater than every x.

Any answer must have two properties, and a different answer can be checked against them.

1. **Every x has a partner:** for each $x \in \mathbb{N}$ there is at least one y with $R(x, y)$.
2. **No y is a partner of every x:** for each $y \in \mathbb{N}$ there is at least one x with $R(x, y)$ false.

For example, "$y \neq x$" also works: each x has the partner $x + 1$, and each y fails to be a partner of $x = y$. So does "$y \geq x$": each x is its own partner, and each y fails to be a partner of $x = y + 1$. A predicate that does **not** work is "$y \geq 1$": property 1 holds, but property 2 fails, because $y = 1$ is a partner of every x, so $\exists y\, \forall x\, (y \geq 1)$ is true.

Exercise. B3: Every and some on three numbers

Let the domain be $D = \{0, 1, 2\}$, with $\leq$ and $<$ as usual, and with the constants $0, 1, 2$ naming its elements.

(a) Write $\exists x\, \forall y\, (x \leq y)$ without quantifiers, as a disjunction of conjunctions, and decide it.

(b) Write $\neg \forall x\, (x < 2)$ without quantifiers, as a disjunction, using De Morgan's law, and

say which disjunct is true.

(c) Let A be the set of elements of D greater than 2. Decide $\forall x \in A\ (x < 0)$ and $\exists x \in A\ (x < 0)$.

(d) Give an example, or explain why there is none: a domain and a one-place predicate P for which $\forall x\, P(x)$ is true and $\exists x\, P(x)$ is false.

Solution. *(a)* The sentence is

$$(0 \le 0 \wedge 0 \le 1 \wedge 0 \le 2) \vee (1 \le 0 \wedge 1 \le 1 \wedge 1 \le 2) \\ \vee (2 \le 0 \wedge 2 \le 1 \wedge 2 \le 2).$$

It is true, by the first disjunct: 0 is a least element.

(b) $\neg(0 < 2 \wedge 1 < 2 \wedge 2 < 2)$ is, by De Morgan's law (Proposition 2.5.5), $\neg(0 < 2) \vee \neg(1 < 2) \vee \neg(2 < 2)$. The third disjunct is true, since $2 < 2$ is false: "some element is not below 2", with the witness 2.

(c) A is empty. The first is true: for every x, the hypothesis $x \in A$ is false. The second is false: no x is in A.

(d) There is none. A domain has at least one element (the convention of "Quantifiers and the domain"), say d; if $\forall x\, P(x)$ is true, then P holds of d, so d is a witness and $\exists x\, P(x)$ is true. The emptiness in (c) is the emptiness of a bound, not of the domain.

Worked examples of these types: the long "and" of world W; "The empty set" in "Bounded quantifiers".

Marking scheme. (10 points)

- 3: (a), all nine atoms in the right places and the verdict with its witness.
- 2: (b), De Morgan applied to a three-term conjunction and the true disjunct named.
- 2: (c), both verdicts, each with the reason (hypothesis false; no element).
- 3: (d), "none", with the argument from a non-empty domain; 1 of the 3 for an answer that separates the empty bound of (c) from the domain.

Fatal error (0 for the part): an empty domain offered as the example in (d).

5.1.9.3 C. Going deeper

Exercise. C1: A least element

Decide whether $\exists x\, \forall y\, (x \le y)$ is true over each of the domains $\mathbb{N}$, $\mathbb{Z}$, $[0, 1]$ and $(0, 1]$. Give the witness where it is true, and where it is false show that every proposed x is defeated by some y.

Solution. The sentence says that the domain has a **least element**: one x that is at most every element y.

Over $\mathbb{N}$: true, with the witness $x = 0$. Every natural number y satisfies $0 \le y$.

Over $\mathbb{Z}$: false. Let x be any integer. The integer $y = x - 1$ satisfies $y < x$, so $x \le y$ fails. No

integer is least.

Over $[0,1]$*: true*, with the witness $x = 0$. Every $y \in [0,1]$ satisfies $0 \leq y$, by the definition of the interval.

Over $(0,1]$*: false.* Let $x \in (0,1]$, so $0 < x \leq 1$. The number $y = \frac{x}{2}$ satisfies $0 < \frac{x}{2} < x \leq 1$, so $y \in (0,1]$ and $y < x$; hence $x \leq y$ fails. The defeating y depends on the proposed x, which is allowed, because in the game the challenger chooses y after seeing x.

The last two domains differ by a single point, 0, and that point is exactly the least element of $[0,1]$. Removing it leaves a set with no least element, although every element of $(0,1]$ is still at least 0. *Analysis on the Real Line*, Chapter 1 returns to this distinction between a least element and a greatest lower bound.

Exercise. C2: The wrong connective

Take the domain $\mathbb{R}$.

(a) The sentence "every element of A is positive" is mistranslated as $\forall x\,(x \in A \wedge x > 0)$, and "some element of A is positive" as $\exists x\,(x \in A \to x > 0)$. For each mistranslation, give a set $A \subseteq \mathbb{R}$ for which it has a different truth value from the English sentence.

(b) Show that the first mistranslation is false for **every** set $A \subseteq \mathbb{R}$, and the second is true for **every** set $A \subseteq \mathbb{R}$. What does this say about how much each one tells us about A?

Solution. (a) Take $A = \{1\}$ for the first. "Every element of $\{1\}$ is positive" is true. The mistranslation claims that every real x is in $\{1\}$ and positive; $x = 0$ is neither, so it is false.

Take $A = \{-1\}$ for the second. "Some element of $\{-1\}$ is positive" is false, since its only element is negative. The mistranslation is true, with the witness $x = 0$: since $0 \notin \{-1\}$, the implication "if $0 \in A$, then $0 > 0$" has a false hypothesis and is true.

(b) For the first, let $A \subseteq \mathbb{R}$ be any set. The real number $x = 0$ does not satisfy $x > 0$, so it does not satisfy $x \in A \wedge x > 0$, whatever A is. It is a counterexample to the $\forall$, so the sentence is false.

For the second, let $A \subseteq \mathbb{R}$ be any set, and consider the real number $x = 1$. The conclusion $1 > 0$ is true, so the implication $1 \in A \to 1 > 0$ is true whether or not $1 \in A$. Hence $x = 1$ is a witness, whatever A is, and the sentence is true.

So neither mistranslation tells us anything about A at all: the first is false and the second is true, whichever set A is. A sentence whose truth value does not depend on A cannot express a property of A. This is the precise sense in which the wrong connective "claims too much" under $\forall$ and "too little" under $\exists$.

Marking scheme. (10 points) - 2: a set for the first mistranslation, with both truth values checked. - 2: a set for the second mistranslation, with both truth values checked, and the implication at the witness justified (its hypothesis is false, or its conclusion is true). - 3: for the first, for an arbitrary A, a real number (the same for every A, or chosen by cases covering every A) that fails $x \in A \wedge x > 0$, with the check. - 3: for the second, for an arbitrary A, a witness (the same for every A, or chosen by cases covering every A) at which $x \in A \to x > 0$ holds, with a reason that covers every A (its conclusion is true, or a check in every case). Fatal errors (0 for the step): a witness or counterexample that depends on A in (b) without

a case split covering every A; reading $\exists x\,(x \in A \to x > 0)$ as "some element of A is positive".

5.2 Terms, Formulas, Free Variables and Substitution

Section 5.1 read quantified sentences by eye. That is enough to understand them, and not enough to reason with them: every quantifier rule of Sections 5.3 to 5.6, and the rule for using an equation in Section 5.7, has the form "put the term t in for the variable x in φ", written $\varphi[t/x]$, and a rule has to say exactly which symbols change. This section makes the language of Section 5.1 exact, in the way Section 4.1 made the propositional language exact, says exactly what $\varphi[t/x]$ does, and names the one way it goes wrong, **capture**.

5.2.1 Which letters are holes

Three sentences, over the domain $\mathbb{N}$:

- "$x > 2$" is not a statement. It is the open sentence of item (e) of Example (Statements and nonstatements), true for some values of x and false for others.
- "$\forall x\,(x^2 \geq 0)$" is a statement, and a true one. The letter x still appears, but the quantifier has closed its hole: the sentence does not ask which x we mean.
- "$\forall x\,(x \geq y)$" is neither true nor false as it stands. It is true when y is 0 (every natural number is at least 0) and false when y is 1 (the number 0 is not at least 1). It is an open sentence **in** y, although it begins with a quantifier.

So the question "is this a statement?" is not answered by looking for quantifiers. It is answered by asking which letters are still holes. In the third sentence x is closed by $\forall x$ and y is open. The open letters will be called **free**, and the closed ones **bound**.

Here is why the difference matters for reasoning. Over $\mathbb{N}$, the sentence

$$\exists y\,(y \neq x)$$

is an open sentence in x, and it is true for every x: given x, the number $y = x + 1$ differs from it. Now suppose that we want to say the same thing about a particular number named by another letter, and we put y in for x without thinking. The result is

$$\exists y\,(y \neq y),$$

"some number differs from itself", which is false. Nothing about numbers went wrong. The y that we put in landed inside the scope of $\exists y$, and the quantifier took hold of it: the new y was meant to name a number from outside, and it has become the quantifier's own placeholder. This is **capture**.

Capture is the one trap of substitution, and the rules of Sections 5.3 to 5.7 substitute at almost every step. So we need exact answers to three questions: what the expressions are that can be put in; which occurrences of x get replaced; and when putting t in for x is safe.

5.2.2 Terms and formulas

The first things to build are the names of objects.

A term is a name of an object, built from variables and constants by function symbols.

Definition 5.2.1. Term

The **terms** are the strings given by the following clauses.

(Tm1) Every variable is a term, and every constant symbol is a term.

(Tm2) If f is a function symbol of arity n and $t_1, \ldots, t_n$ are terms, then $f(t_1, \ldots, t_n)$ is a term.

(Tm3) **Nothing else** is a term: a string is a term **only if** it is the last entry of a finite list of strings in which every entry is a variable, a constant symbol, or $f(s_1, \ldots, s_n)$ for some function symbol f of arity n and **earlier** entries $s_1, \ldots, s_n$.

In words: variables and constants are terms, a function symbol applied to as many terms as it has places is a term, and nothing else is.

A **variable of** a term t is a variable that occurs in t. A term with no variables, such as 2 or $f(c)$, names one fixed object once the symbols are interpreted; a term with variables, such as $x + 1$, names an object only once its variables have values.

Shorthand. We write terms in the familiar way: $x + y$ for $+(x, y)$, x^2, $|x|$, a_n for $a(n)$ (a sequence is a function symbol, as Section 5.1 said). School conventions about grouping say which term is meant.

Here are terms, simplest first.

- x and 2: by (Tm1). The degenerate terms, built in no steps; every argument about all terms has to handle them.
- $x + 1$, officially $+(x, 1)$: by (Tm2) from the terms x and 1.
- $f(f(c))$: by (Tm2) twice, from c, for a one-place f. Its only building blocks are constants, so it has no variables.
- $|a_n - L|$, officially $|\cdot|(-(a(n), L))$, with the two-place minus: by (Tm2) three times, from the variables n and L. Its variables are n and L. The letter a is a function symbol, not a variable.

↻ **Second reading** · *The symbols of the language, strings that are not terms, and two more shorthand rules*

Section 4.1 described the propositional language by its alphabet and its formation rules. The language of predicate logic has more symbols, of more kinds.

- **Variables:** x, y, z, u, with subscripts $x_0, x_1, y_0, \ldots$, and in analysis also $\varepsilon, \delta, N, n, k, L$. There are infinitely many, so that a fresh one is always available. Each is one symbol.
- **Constant symbols**, such as 0, 1, 2, a, b, c. (A letter may be a constant in one language and a function symbol in another: in a language with a sequence, a is the sequence's one-place function symbol, as below, and not a constant.)
- **Function symbols**, each with a number of places $n \geq 1$, its **arity**: f and g, the arithmetic symbols $+$, $\cdot$ and $-$ (two places), and $-$ in front of a number, halving $\cdot/2$, squaring and $|\cdot|$ (one place). The one-place and the two-place minus are different symbols written alike; the number of terms they are applied to shows which is meant.
- **Predicate symbols**, each with an arity $n \geq 1$: P, Q, R, S, $<$, $\leq$, $>$, $\geq$, Prime, and $\in A$ for a fixed set A (one place, as in Section 5.1).
- **Equality** $=$, which every language has.
- The constant $\bot$, the connectives $\neg, \wedge, \vee, \to, \leftrightarrow$ of Chapter 4, and the **quantifier symbols**

$\forall$ and $\exists$.
- Parentheses (,), and the comma.

Which constant, function, and predicate symbols are available depends on the subject. For the argument of Section 5.1 we need the constant 2 and the one-place predicate symbols P and S. For arithmetic we need 0, 1, $+$, $\cdot$, and $<$. Every formula uses only finitely many symbols, and we never need to fix the whole list in advance.

Strings that are **not** terms, each by one clause:

- $x < 1$ is not a term. It does not name an object; it says something about one. It is built with a predicate symbol, and (Tm2) uses only function symbols. It will be a formula.
- $f(x, y)$, for a one-place f, is not a term: (Tm2) applies f to exactly one term.

Terms can be drawn as trees, exactly as formulas were in Section 4.1: a variable or a constant is a single node, and $f(t_1, \ldots, t_n)$ is a node labeled f with the trees of $t_1, \ldots, t_n$ below it, in order. And a term can be read in only one way. The proof is the proof of unique readability for formulas in Section 4.3 (Theorem 4.3.3), with the parentheses and commas of (Tm2) doing the work that the parentheses of (F3) did there. We do not repeat it.

Shorthand rules 1 and 3.

Rule 1. Two-place predicate symbols such as $<, \leq$ are written between their terms: $x < y$ is shorthand for $<(x, y)$. The one-place symbol $\in A$ is written after its term: $t \in A$ is shorthand for $(\in A)(t)$. (When sets are themselves objects of the domain, $\in$ is a two-place predicate symbol and $x \in A$ is $\in (x, A)$; in this chapter A is a fixed set, except in one example of Section 5.7 where sets are the objects, so we treat $\in A$ as one-place.) We write $t_1 \neq t_2$ for $\neg(t_1 = t_2)$. **Rule 3.** The bounded quantifiers of Definition 5.1.1, $\forall x \in A\ \varphi$ and the rest, stand for the formulas given there.

Next, the statements and open sentences. Atomic formulas say something about terms; everything else is built from them by connectives and quantifiers.

A first-order formula is built from predicates applied to terms, by the connectives and the two quantifiers.

Definition 5.2.2. First-Order Formula

The **formulas** (of predicate logic, or **first-order formulas**) are the strings given by the following clauses.

(Fm1) **Atomic formulas.** If R is a predicate symbol of arity n and $t_1, \ldots, t_n$ are terms, then $R(t_1, \ldots, t_n)$ is a formula. If t_1 and t_2 are terms, then $t_1 = t_2$ is a formula. And $\bot$ is a formula.

(Fm2) If φ is a formula, then $\neg\varphi$ is a formula.

(Fm3) If φ and ψ are formulas, then $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \to \psi)$, and $(\varphi \leftrightarrow \psi)$ are formulas.

(Fm4) If φ is a formula and x is a **variable**, then $\forall x\, \varphi$ and $\exists x\, \varphi$ are formulas.

(Fm5) **Nothing else** is a formula: a string is a formula only if it is the last entry of a finite list in which every entry is an atomic formula, or is built by (Fm2), (Fm3) or (Fm4) from **earlier** entries.

- *(Fm1)*: the atomic formulas, a predicate symbol with its terms, or an equation, or $\bot$, take the place of Chapter 4's atoms.
- *(Fm2)*, *(Fm3)*: the connectives, as before.
- *(Fm4)*: puts $\forall x$ or $\exists x$, with a **variable** x, in front of a formula, without parentheses, as $\neg$ is.

Structural induction and recursion (Theorem 4.3.1, Theorem 4.3.4) carry over to terms and first-order formulas, with the atomic formulas as base case and one new building step for each quantifier (the fold below says how).

↻ **Second reading** · *Induction and recursion on formulas*

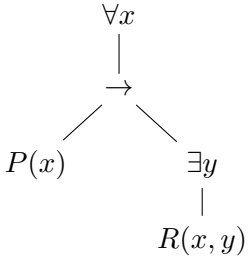
The base case covers every **atomic** formula (and $\bot$), and there are two new building steps, from φ to $\forall x\,\varphi$ and to $\exists x\,\varphi$, for every variable x, each with one part, like the step to $\neg\varphi$. So, to prove that every formula has a property, prove it for the atomic formulas and $\bot$, and prove that it passes from φ to $\neg\varphi$, $\forall x\,\varphi$, and $\exists x\,\varphi$, and from φ, ψ to $(\varphi \circ \psi)$. To define a function on formulas, give one clause for each of these cases. The proofs in Section 4.3 go through word for word with the new cases added, and unique readability holds for the same reason; we use all three without repeating them. The same two principles hold for terms, with (Tm1) as the base case and (Tm2) as the one building step.

Shorthand. The conventions of Section 4.1 for dropping parentheses stay, with three additions. Rules 1 and 3 (predicate symbols written between their terms, as in $x < y$; bounded quantifiers as abbreviations) are in the fold on the symbols of the language, above; rule 2 matters in every section:

Rule 2. A quantifier binds as tightly as $\neg$: it applies to the smallest formula that follows it. So $\forall x\,P(x) \to Q(c)$ stands for $(\forall x\,P(x) \to Q(c))$, an implication whose hypothesis is $\forall x\,P(x)$. For the quantifier to govern the whole implication, write $\forall x\,(P(x) \to Q(x))$, where the parentheses are the official ones around the implication.

Rule 2 matters in every rule section. The formula $\forall x\,P(x) \to Q(c)$ has main connective $\to$, while $\forall x\,(P(x) \to Q(x))$ has main symbol $\forall x$. The rule for using a "for every" (Section 5.3) applies only to the second.

Parse trees. A parse tree for a first-order formula is drawn as in Section 4.1, with one new kind of node: a **quantifier node**, labeled $\forall x$ or $\exists x$, with the tree of the formula it governs as its only subtree. The leaves are atomic formulas (or $\bot$); their terms have trees of their own, which we do not draw. Here is the tree of $\forall x\,(P(x) \to \exists y\,R(x,y))$.



The root is the quantifier node $\forall x$, so the formula's main symbol is $\forall x$. Below it is the implication, with the atomic formula $P(x)$ on the left and the quantifier node $\exists y$ on the right, which

governs the atomic formula $R(x, y)$.

Quick check. Draw the parse tree of $\forall x\, P(x) \to \exists y \forall x\, R(x, y)$, written in shorthand. What is its main symbol?

Solution. By shorthand rule 2, each quantifier applies to the smallest formula after it, so the official formula is $(\forall x\, P(x) \to \exists y \forall x\, R(x, y))$ and its main symbol is $\to$. The root is $\to$. Its left subtree is the node $\forall x$ above the leaf $P(x)$. Its right subtree is the node $\exists y$, below it the node $\forall x$, and below that the leaf $R(x, y)$. The two $\forall x$ nodes are on different branches.

5.2.3 Free and bound: drawing the arrows

The parse tree shows which quantifier governs which part of a formula. That is exactly what decides whether a letter is a hole. Draw an arrow from each occurrence of a variable to the nearest quantifier on that variable above it in the parse tree. An occurrence with an arrow is bound; one with none is a hole still open.

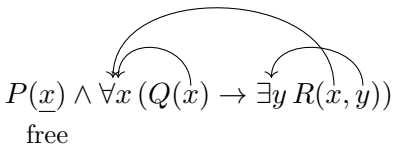


Figure B. Each bound occurrence points to its quantifier. The first x points nowhere: it is free.

An occurrence of a variable is bound if a quantifier on that variable stands above it in the parse tree; otherwise it is free.

Definition 5.2.3. Free and Bound Variables

Let φ be a formula and x a variable. An **occurrence** of x in φ is an appearance of x inside one of the terms of an atomic formula of φ. (The x written just after a quantifier symbol, as in $\forall x$, is part of the quantifier node and is not an occurrence.)

- An occurrence of x is **bound** if the path from its leaf up to the root of the parse tree passes through a quantifier node $\forall x$ or $\exists x$, with the **same** variable x. The lowest such node **binds** the occurrence.
- An occurrence that is not bound is **free**.
- The **scope** of a quantifier node is the part of the tree below it: the formula it governs.
- The variable x is a **free variable of** φ, or is **free in** φ, if at least one occurrence of x in φ is free.

A variable **appears in** φ if it is written anywhere in φ: inside a term, or just after a quantifier symbol. Every freshness condition of this chapter is stated with "appear"; in practice, a fresh letter is one written nowhere yet.

Definition 5.2.4. Sentence

A **sentence** is a formula with no free variables.

A sentence is what Chapter 1 called a statement, once the domain and the symbols are fixed: every hole is closed, so it is true or false. A formula with free variables is an open sentence, true or false only once values are given to its free variables.

Examples. The domain is irrelevant to all of them: being free or bound is a matter of the string, not of what it means.

- $\forall x\,(x \geq y)$: the occurrence of x is bound by $\forall x$; the occurrence of y has no quantifier on y above it, so it is free. The free variables are y alone. Not a sentence.
- $P(x) \wedge \forall x\, Q(x)$: the occurrence of x in $P(x)$ is **free** (no quantifier stands above it: the $\forall x$ governs only $Q(x)$), and the occurrence in $Q(x)$ is bound. So x is a free variable of this formula, and it is not a sentence.
- $\forall x\, \exists x\, P(x)$: the occurrence of x has two quantifiers on x above it, and the lower one, $\exists x$, binds it. The outer $\forall x$ binds nothing. A sentence.

↻ **Second reading** · *Occurrences, 'appear', and the recursive description of 'free'*

In words: an occurrence of x is bound when it lies in the scope of a quantifier on x; quantifiers on other variables do not bind it. If several quantifiers on x stand above it, the nearest one, the lowest in the tree, is the one that binds it. A free occurrence is a hole still open. A variable is free in a formula if it has **at least one** free occurrence, even if it has bound occurrences too.

So y appears in $\forall y\, P(x)$, although it has no occurrence there. A variable that does not appear in φ has no occurrence in φ, and no quantifier of φ is on it; that is why the freshness conditions use "appear".

Three more examples:

- $\bot$ and $P(c)$ are sentences: they contain no variables at all. So is $\forall x\, P(c)$. These are the degenerate cases.
- $\forall x\,(x^2 \geq 1)$: the single occurrence of x lies below the node $\forall x$, so it is bound. A sentence.
- $\forall x\,\big(P(x) \to \exists y\, R(x,y)\big)$, the parse tree drawn above: both occurrences of x are bound by the root; the occurrence of y in $R(x,y)$ is bound by $\exists y$. A sentence.

The recursive description. In proofs about all formulas we need "x is free in φ" in the form that structural induction can use, one case per clause of Definition 5.2.2.

- x is free in an atomic formula exactly when x occurs in one of its terms; x is never free in $\bot$.
- x is free in $\neg\varphi$ exactly when x is free in φ.
- x is free in $(\varphi \circ \psi)$ exactly when x is free in φ or in ψ.
- x is free in $\forall y\, \varphi$, and in $\exists y\, \varphi$, exactly when x is a different variable from y **and** x is free in φ.

Each line is read off the tree. In an atomic formula there are no quantifier nodes, so every occurrence is free. The tree of $\neg\varphi$ or $(\varphi \circ \psi)$ adds a root that is not a quantifier node, so it binds nothing. The tree of $\forall y\, \varphi$ adds the root $\forall y$ on top of the tree of φ: an occurrence of x inside φ keeps its path, with one more node at the top, so it is free in $\forall y\, \varphi$ exactly when it was free in φ

and the new node is not a quantifier on x, that is, when x and y are different variables.

↻ **Second reading** · *The free variables of a limit statement*

Example. The free variables of a limit statement

With the conventions of Section 5.1 (domain $\mathbb{R}$, the sequence as a function symbol a), "the sequence (a_n) converges to L" is the formula

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n \in \mathbb{N} \; (n \geq N \to |a_n - L| < \varepsilon).$$

Write it without bounded quantifiers, and find its free variables. Is it a sentence?

Solution. Unfolding the three bounded quantifiers by Definition 5.1.1 ($\to$ under each $\forall$, $\wedge$ under the $\exists$) gives

$$\forall \varepsilon \Big(\varepsilon > 0 \to \exists N \, (N \in \mathbb{N} \wedge \forall n \, (n \in \mathbb{N} \to (n \geq N \to |a_n - L| < \varepsilon))) \Big).$$

The variables that occur are ε, N, n, and L. Check each.

- Every occurrence of ε lies below the root $\forall \varepsilon$: bound.
- Every occurrence of N lies below $\exists N$: bound.
- Every occurrence of n (in $n \in \mathbb{N}$, in $n \geq N$, and in the term a_n) lies below $\forall n$: bound.
- The occurrence of L has no quantifier on L above it: free.

So L is the only free variable, and the formula is **not** a sentence. It is an open sentence in L, as it should be: until L is given a value, "(a_n) converges to L" is neither true nor false, and which values make it true depends on the sequence (for $a_n = 1/(n+1)$ the value $L = 0$ does and $L = 1$ does not; for $a_n = (-1)^n$ no value does). Putting $\exists L$ in front gives the sentence "(a_n) converges".

The other symbols in it are not variables, and "free" and "bound" do not apply to them. The letter a is a function symbol; the formula is certainly about the sequence a, but a is fixed by the language, like $+$, and cannot be quantified in this logic. The symbol $\in \mathbb{N}$ is a one-place predicate symbol, 0 is a constant, and $>$, $\geq$, $<$, $-$ and $|\cdot|$ are predicate and function symbols of the language.

Warning. The same letter can occur free and bound in one formula. In $P(x) \wedge \forall x \, Q(x)$, the first x is a hole and the second is the quantifier's placeholder, and the two have nothing to do with each other. This is legal, and it is a trap for the reader: a proof that "uses x" from such a formula has to say which x it means. The book avoids it by renaming the bound variable (the last part of this section explains how): $P(x) \wedge \forall z \, Q(z)$ says the same thing, and the question does not arise.

Quick check. Is x free in $\forall y\,(x < y) \to \exists x\,(x = 0)$? At which occurrences?

Solution. Yes. By shorthand rule 2 the formula is $(\forall y\,(x < y) \to \exists x\,(x = 0))$. The occurrence of x in $x < y$ lies below $\forall y$ only, which is a quantifier on y, so it is **free**. The occurrence of x in $x = 0$ lies below $\exists x$, so it is bound. One free occurrence is enough: x is a free variable of the formula. The variable y, by contrast, is not free: its only occurrence lies below $\forall y$.

5.2.4 Substitution, by example first

We now come to the operation that every quantifier rule uses. The rule of Section 5.3 will take the premise $\forall x\,(P(x) \to S(x))$ of Section 5.1's argument and produce $P(2) \to S(2)$: it strips the quantifier and puts 2 in for x. The rule of Section 5.5 goes the other way. Both need to say exactly what "put 2 in for x" does. The answer has one principle: only the **free** occurrences are replaced, because only they are holes. A bound occurrence belongs to its quantifier, and the quantifier is not asking for a value.

$\varphi[t/x]$ is φ with t put in for every free occurrence of x, and nothing else changed.

Example. Four substitutions

Compute each substitution by the rule in the slogan above.

(a) $(P(x) \to S(x))[2/x]$.
(b) $(x < y \wedge \forall x\,(x \geq 0))[(y+1)/x]$.
(c) $(\forall y\,R(x,y))[f(z)/x]$.
(d) $(\forall x\,P(x))[c/x]$.

Solution. (a) The implication passes the substitution to both sides, and in each atomic formula the term x becomes 2: the result is $P(2) \to S(2)$. This is the step that Section 5.3's rule makes in the argument of Section 5.1.

(b) The conjunction passes it to both parts. In $x < y$ the x becomes $y+1$, giving $y+1 < y$. In $\forall x\,(x \geq 0)$ the quantifier is on x, so the substitution stops. The result is

$$y + 1 < y \ \wedge\ \forall x\,(x \geq 0).$$

Only the free occurrence of x changed. The bound one, in $x \geq 0$, was the quantifier's placeholder, and replacing it would have turned the sentence "every x is at least 0" into a claim about $y+1$.

(c) The quantifier is on y, a variable different from x, so it passes the substitution on: the result is $\forall y\,R(f(z), y)$.

(d) The quantifier is on x, so nothing changes: the result is $\forall x\,P(x)$. There was no free x to replace.

Here are the clauses those computations followed.

Definition 5.2.5. Substitution of a Term for a Variable

Let t be a term and x a variable. For a term s, the term $s[t/x]$ is s with every occurrence of x replaced by t; by recursion on terms,

$$x[t/x] = t, \qquad y[t/x] = y \ \text{ for a variable } y \neq x, \qquad c[t/x] = c,$$
$$f(s_1, \ldots, s_n)[t/x] = f(s_1[t/x], \ldots, s_n[t/x]).$$

For a formula φ, the formula $\varphi[t/x]$, read "φ with t for x", is defined by recursion on φ:

$$\begin{aligned} R(s_1, \ldots, s_n)[t/x] &= R(s_1[t/x], \ldots, s_n[t/x]), \\ (s_1 = s_2)[t/x] &= (s_1[t/x] = s_2[t/x]), \qquad \bot[t/x] = \bot, \\ (\neg\varphi)[t/x] &= \neg(\varphi[t/x]), \\ (\varphi \circ \psi)[t/x] &= (\varphi[t/x] \circ \psi[t/x]), \\ (\forall x\, \varphi)[t/x] &= \forall x\, \varphi, \\ (\forall y\, \varphi)[t/x] &= \forall y\, (\varphi[t/x]) \ \text{ for a variable } y \neq x, \end{aligned}$$

and the last two lines likewise with $\exists$ in place of $\forall$.

A compound term is written in parentheses: $\varphi[(y+1)/x]$ puts $y+1$ in for x.

Two facts follow from the definition at once, and both are used again and again.

First, $\varphi[x/x] = \varphi$: putting x in for itself changes nothing. In each clause the replaced symbol is x again.

Second, and more useful: **if x is not free in φ, then $\varphi[t/x] = \varphi$**, for every term t. Example (d) was one instance. Later sections use it; its proof, a structural induction, is in the fold below.

Second reading · *The clauses in words, why the definition defines, and the claim proved*

In words, clause by clause: in a term, every x is replaced, since a term has no quantifiers. In an atomic formula, the replacement happens inside its terms. The connectives pass the substitution to their parts. A quantifier on x **stops** it: below $\forall x$ there is no free occurrence of x to replace. A quantifier on another variable y passes it on.

Well defined. The clauses give one value for each case of Definition 5.2.2, so structural recursion (carried over as above) says that they define exactly one formula $\varphi[t/x]$ for each φ. And the result is a formula: each clause builds its value by the clauses of Definition 5.2.1 and Definition 5.2.2 from values already known to be terms and formulas.

This is a different operation from the substitution σ of Section 4.5 (the one in Lemma 4.5.3). There, formulas were put in for **atoms**, at every occurrence. Here a **term** is put in for a **variable**, and bound occurrences are left alone.

Claim. Let x be a variable and t a term. For every formula φ: if x is not free in φ, then $\varphi[t/x] = \varphi$.

Idea. Substitution changes only free occurrences of x, and there are none. Formally, structural induction on φ, using the recursive description of "free" case by case. The only case with any content is a quantifier on a variable other than x, where the substitution is passed on, and the induction hypothesis applies because x is still not free inside.

Proof. For terms first: if x does not occur in a term s, then $s[t/x] = s$, by induction on s (a variable other than x and a constant are unchanged by the clauses, and $f(s_1, \dots, s_n)[t/x] = f(s_1, \dots, s_n)$ when each $s_i[t/x] = s_i$). Now we prove the claim by structural induction on φ.

Atomic formulas and $\bot$. If x is not free in $R(s_1, \dots, s_n)$, then x occurs in none of the s_i (the recursive description), so each $s_i[t/x] = s_i$, and the clause for atomic formulas returns $R(s_1, \dots, s_n)$. The same holds for $s_1 = s_2$, and $\bot[t/x] = \bot$.

Connectives. If x is not free in $\neg\varphi$, it is not free in φ, so $(\neg\varphi)[t/x] = \neg(\varphi[t/x]) = \neg\varphi$ by the induction hypothesis. If x is not free in $(\varphi \circ \psi)$, it is free in neither part, and the same argument applies to both.

Quantifiers. Consider $\forall y\, \varphi$. If y is x, then $(\forall x\, \varphi)[t/x] = \forall x\, \varphi$ by the definition, whatever φ is. If y is a different variable and x is not free in $\forall y\, \varphi$, then x is not free in φ (the recursive description), so $(\forall y\, \varphi)[t/x] = \forall y\, (\varphi[t/x]) = \forall y\, \varphi$ by the induction hypothesis. The case of $\exists y\, \varphi$ is the same.

This proves the claim for every formula φ.

■

One substitution at a time. $\varphi[t/x]$ replaces one variable. Two substitutions in a row are made one after the other, and the second acts on the result of the first. So $R(x, y)[y/x]$ is $R(y, y)$, and then $[x/y]$ turns it into $R(x, x)$, not into the swapped $R(y, x)$: the first step has already lost track of which y was which. To swap two variables, go through a letter that appears nowhere: $[z/x]$, $[x/y]$, $[y/z]$ give $R(z, y)$, $R(z, x)$, $R(y, x)$.

Quick check. Compute $((P(x) \wedge Q(y))[y/x])[x/y]$. Which formula would swapping x and y give?

Solution. The first step gives $P(y) \wedge Q(y)$, and the second $P(x) \wedge Q(x)$. The swap is $P(y) \wedge Q(x)$, reached by $[z/x]$, $[x/y]$, $[y/z]$ with a new letter z.

5.2.5 Capture, and the practical rule

The definition of substitution never refuses. Given any φ, t and x, it returns a formula. Sometimes that formula does not say what we meant.

Take $\varphi = \exists y\, (y \neq x)$, the example at the start of the section, and $t = y$. The quantifier $\exists y$ is on

a variable different from x, so the definition passes the substitution through, and

$$(\exists y\,(y \neq x))[y/x] = \exists y\,(y \neq y).$$

The definition was followed correctly. The meaning is wrong. The formula $\exists y\,(y \neq x)$ says something about the number x: some number differs from it, which over $\mathbb{N}$ is true whatever x is. We wanted to say the same thing about the number y, and we got "some number differs from itself", which is false. The y we put in was supposed to name a number from outside, and it landed in the scope of $\exists y$ and was **captured**.

Compare a substitution that is safe. Over $\mathbb{R}$, $\forall z\,(x < z)$ says that x is smaller than every number (false, as it happens, for every x; but what matters is what it says). Put $y + 1$ in for x: the quantifier $\forall z$ passes the substitution on, and

$$(\forall z\,(x < z))[(y + 1)/x] = \forall z\,(y + 1 < z),$$

which says exactly that $y + 1$ is smaller than every number. Nothing was captured, because the term $y + 1$ does not contain z, the only variable quantified above the occurrence of x.

The difference between the safe case and the unsafe one is exactly this: whether a variable of t is quantified somewhere above a free occurrence of x.

A term is free for x when putting it in for x lets none of its variables be captured.

Definition 5.2.6. Free For

Let t be a term, x a variable, and φ a formula. The term t is **free for** x in φ if **no** free occurrence of x in φ lies in the scope of a quantifier $\forall y$ or $\exists y$ such that y is a variable of t.

In words: look at each place where t will go, that is, at each free occurrence of x. Look at the quantifiers above it in the tree. If none of them is on a variable of t, the substitution is safe there. The term is free for x when it is safe at every such place. Informally: **no variable of t becomes bound** when t is put in.

Four special cases, each immediate from the definition, cover most uses.

- A term with **no variables** (2, c, $f(c)$) is free for every variable in every formula: there is no y to capture.
- The variable x is free for x in every formula. A free occurrence of x never lies in the scope of a quantifier on x; if it did, it would be bound.
- A variable z that does **not appear** in φ is free for every variable in φ: no quantifier of φ is on z.
- If x is **not free** in φ, every term is free for x in φ: there are no free occurrences to check. (And then $\varphi[t/x] = \varphi$ by the claim above.)

In our examples: y is **not** free for x in $\exists y\,(y \neq x)$, because the free occurrence of x lies in the scope of $\exists y$ and y is a variable of the term y. Likewise $y + 1$ is not free for x in $\exists y\,(y > x)$. And $y + 1$ **is** free for x in $\forall z\,(x < z)$: the only quantifier above the free occurrence of x is on z, which is not a variable of $y + 1$.

"Free for" is the precise condition under which $\varphi[t/x]$ says about t what φ says about x. The rules $\forall$E (Section 5.3), $\exists$I (Section 5.5) and $=$E (Section 5.7) carry it as a side condition.

it needs

Here is a second example, with a compound term. Over $\mathbb{R}$, $\exists y\,(y > x)$ is true for every x: there is a number larger than x. Put $y + 1$ in for x:

$$(\exists y\,(y > x))[(y + 1)/x] = \exists y\,(y > y + 1).$$

The result says that some number is larger than its own successor, which is false. It should have said that some number is larger than $y + 1$, which is true.

The recursive description. For induction proofs, "free for" can be read off case by case, like "free":

- t is free for x in every atomic formula and in $\bot$;
- t is free for x in $\neg\varphi$ exactly when it is free for x in φ, and in $(\varphi \circ \psi)$ exactly when it is free for x in both φ and ψ;
- t is free for x in $\forall y\,\varphi$ (and in $\exists y\,\varphi$) exactly when **either** x is not free in $\forall y\,\varphi$, **or** y is not a variable of t and t is free for x in φ.

The first two lines hold because an atomic formula has no quantifiers and a connective node is not a quantifier. For the third: if x is not free in $\forall y\,\varphi$, there is nothing to check. If it is free, then y is a different variable and the free occurrences of x in $\forall y\,\varphi$ are those in φ; each lies in the scope of the new root $\forall y$ and of the quantifiers inside φ above it. The root is harmless exactly when y is not a variable of t, and the quantifiers inside are harmless exactly when t is free for x in φ.

Section 5.3 shows the capture example of this section turning a true premise into a false conclusion when the side condition is ignored. The definition asks for no more than it needs: a variable of t may be quantified elsewhere in φ, as long as no free occurrence of x lies in its scope. For instance, y is free for x in $P(x) \wedge \exists y\,Q(y)$, and the result $P(y) \wedge \exists y\,Q(y)$ says of y what the original said of x.

Before you substitute. Look at each free occurrence of x and at the quantifiers above it. If none of them is on a letter of t, go ahead. If one is, rename that quantifier's letter to a brand-new letter first (next subsection), and then substitute. Nothing else can go wrong.

Quick check. Is y free for x in $\exists y\,(x < y) \wedge P(y)$? Is $x + 1$?

Solution. By shorthand rule 2 the formula is $(\exists y\,(x < y) \wedge P(y))$. The term y is not free for x: the free x of $x < y$ lies under $\exists y$, and y is a letter of the term. Rename first, $\exists z\,(x < z) \wedge P(y)$, and then substitute to get $\exists z\,(y < z) \wedge P(y)$. The term $x + 1$ is free for x: the only quantifier above the free x is on y, which is not a letter of $x + 1$.

Warning. Capture changes meaning, and it gives no warning: the definition of $\varphi[t/x]$ returns a formula either way. Over $\mathbb{R}$, $\exists y\,(y > x)$ is true for every x, and $(\exists y\,(y > x))[(y + 1)/x]$ is $\exists y\,(y > y + 1)$, which is false. Before putting t in for x, check that no variable of t is quantified above a free occurrence of x.

Quick check. Is $y+1$ free for x in $\forall z\,(x < z)$? Is it free for x in $\forall y\,(x < y)$? Compute both substitutions and say what each result claims.

Solution. In $\forall z\,(x < z)$: yes. The only quantifier above the occurrence of x is on z, and z is not a variable of $y+1$. The result is $\forall z\,(y+1 < z)$, "$y+1$ is smaller than every number", which is what the original said about x.

In $\forall y\,(x < y)$: no. The occurrence of x lies in the scope of $\forall y$, and y is a variable of $y+1$. The definition gives $\forall y\,(y+1 < y)$: "every number is larger than its own successor". That is a sentence about no particular number at all, and a false one; the original said something about x, and the result should have said the same about $y+1$. The y of $y+1$ was captured.

5.2.6 Renaming bound variables

What should we do when t is not free for x in φ? A bound variable is a placeholder, and the name of a placeholder does not matter. "Some number differs from x" can be written $\exists y\,(y \neq x)$ or $\exists z\,(z \neq x)$, just as $\int_0^1 s^2\,ds$ and $\int_0^1 u^2\,du$ are the same integral. So the way out is: **rename the bound variable first, then substitute.**

Renaming. Let φ be a formula and y, z variables, with z **not appearing** in φ. The formula $\forall z\,\varphi[z/y]$ is obtained from $\forall y\,\varphi$ by **renaming** the bound variable y to z; likewise $\exists z\,\varphi[z/y]$ from $\exists y\,\varphi$. The same may be done to a quantifier anywhere inside a larger formula.

In words: change the variable after the quantifier, and change the occurrences it binds, to a letter that appears nowhere in the formula. The condition that z does not appear in φ does two jobs. It makes z free for y in φ (the third special case above), so no new z is captured by a quantifier inside φ. And it ensures that no free z already in φ is captured by the new quantifier $\forall z$. Without it, renaming can change the meaning: $\forall y\,(y \geq z)$ says something about z (over $\mathbb{N}$, that $z = 0$), and "renaming" y to z would give $\forall z\,(z \geq z)$, a sentence that is simply true.

Renaming a bound letter is treated as never changing what a formula says; Sections 5.4 and 5.6 derive it at the front of a formula, and the fold says where the rest stands.

↻ **Second reading** · *The status of renaming*

Its status. Renaming is not a rule of the calculus, and we have not proved that it keeps meaning. Here is where it stands. For a renaming at the front of a formula, $\forall y\,\varphi$ and $\forall z\,\varphi[z/y]$ can each be derived from the other in three lines, once the rules of Section 5.4 are available; that section does it, and Section 5.6 does the same for $\exists$. A renaming **inside** a larger formula we take as a convention of this book: formulas that differ only by renaming bound variables are treated as saying the same thing. Its justification in terms of meaning is the one the Looking-ahead box at the end of this section points to.

The recipe. To put t in for x in φ when t is not free for x: rename each quantifier of φ that is on a variable of t and has a free occurrence of x in its scope, each to a different variable that appears neither in φ nor in t (each such step is a renaming in the sense above); then substitute.

After the renaming, t is free for x.

- $\exists y\,(y \neq x)$ with y for x: rename y to z, giving $\exists z\,(z \neq x)$; then substitute, giving $\exists z\,(z \neq y)$, "some number differs from y". True over $\mathbb{N}$ for every y, as intended.
- $\exists y\,(y > x)$ with $y + 1$ for x: rename to $\exists z\,(z > x)$; substitute, giving $\exists z\,(z > y + 1)$. True over $\mathbb{R}$, as intended.

Which fresh letter we choose does not matter: two choices give formulas that differ only by a renaming. From now on, whenever a rule writes $\varphi[t/x]$, it will require t to be free for x in φ, and if we want a t that is not, we rename first.

5.2.7 How substitution looks in a mathematical proof

Mathematicians substitute constantly and never write $[t/x]$. Two examples show the operation, and its trap, in their natural habitat.

"Apply the hypothesis with $\varepsilon/2$ in place of ε." Suppose we know that (a_n) converges to L, and we have fixed an $\varepsilon > 0$ for some other purpose. A proof says: "applying the hypothesis with $\varepsilon/2$, there is an $N \in \mathbb{N}$ with $|a_n - L| < \varepsilon/2$ for all $n \geq N$". The hypothesis is $\forall\varepsilon\,\varphi$, with

$$\varphi \;=\; \Big(\varepsilon > 0 \to \exists N \in \mathbb{N}\,\forall n \in \mathbb{N}\,\big(n \geq N \to |a_n - L| < \varepsilon\big)\Big),$$

and the proof uses the instance $\varphi[t/\varepsilon]$ with the term $t = \varepsilon/2$:

$$\varepsilon/2 > 0 \to \exists N \in \mathbb{N}\,\forall n \in \mathbb{N}\,\big(n \geq N \to |a_n - L| < \varepsilon/2\big).$$

The term $\varepsilon/2$ contains the variable ε: the fixed ε of the proof, not the bound one of the hypothesis. It is free for ε in φ, because the only quantifiers above the free occurrences of ε in φ are on N and n. Section 5.3's rule makes this step, and "$\varepsilon/2 > 0$" then comes from $\varepsilon > 0$ by arithmetic.

"Take $N = n$." Now suppose the proof has chosen N, so it knows

$$\psi \;=\; \forall n \in \mathbb{N}\,\big(n \geq N \to |a_n - L| < \varepsilon\big),$$

with N free. A tempting step is to "apply this with n in place of N". No rule allows this: N has been chosen, not quantified, so no term may be put in for it. Writing the step out shows how badly it goes wrong, because n is not even free for N in ψ: the occurrence of N lies in the scope of $\forall n$. The substitution produces

$$\forall n \in \mathbb{N}\,\big(n \geq n \to |a_n - L| < \varepsilon\big),$$

which says that **every** term of the sequence is within ε of L, a far stronger claim that does not follow. In words, the error is to let the threshold N be "the n we are looking at", when n is the placeholder of "for all $n \geq N$". A proof that needs a statement about a particular index names it with a new letter, say $m \geq N$, and substitutes m, which is free for n.

↻ **Second reading** · *Induction, as a formula*

Induction, as a formula. The display after Axiom 4.2.1 in Section 4.2 used a property $P(n)$. With substitution it can be written for any formula φ in which the variable n is free, as the formula

$$\Big(\varphi[0/n] \;\wedge\; \forall n \in \mathbb{N}\,\big(\varphi \to \varphi[(n+1)/n]\big)\Big) \;\to\; \forall n \in \mathbb{N}\,\varphi,$$

where $\varphi[0/n]$ plays the part of $P(0)$ and $\varphi[(n+1)/n]$ the part of $P(n+1)$. The term $n+1$ is free for n in φ: its only variable is n, and a free occurrence of n never lies in the scope of a quantifier on n. This is how the principle of induction reads in the language of this chapter; its status (a basic property of $\mathbb{N}$) is unchanged.

Second reading · *Partial terms*

Mathematics writes $1/x$ and $\lim a_n$, which name nothing for some values: $x = 0$, or a sequence with no limit. In the language of this chapter every term names an object, because every function symbol is defined everywhere on the domain. Mathematics keeps partial operations out of trouble with hypotheses: $1/x$ is written only under $x \neq 0$, as in $\forall x\,(x \neq 0 \to x \cdot (1/x) = 1)$, where the value at 0 is never used; and "the limit" is written only after it has been proved that there is exactly one (Section 5.8). Some logics allow terms that name nothing; this book does not need them.

Remark (Looking ahead). Chapter 7 proves, in an optional section, that renaming a bound letter never changes the meaning of a formula.

5.2.8 Exercises

5.2.8.1 A. Check your understanding

Exercise. A1

(a) For each formula, mark each occurrence of each variable as free or bound, list the free variables, and say whether it is a sentence: (i) $\forall x\,(P(x) \to Q(y))$; (ii) $\exists x\,\forall y\,R(x,y)$; (iii) $\forall x\,(x < y) \wedge \exists y\,(y < x)$; (iv) $\forall \varepsilon\,(\varepsilon > 0 \to \exists \delta\,(\delta > 0 \wedge \delta < \varepsilon))$.

(b) Which of the following are terms, and which are formulas, for one-place f, two-place g and two-place predicate R? (i) $g(f(x), c)$; (ii) $R(f(x), c)$; (iii) $f(R(x, c))$; (iv) $\forall c\,R(c, c)$.

(c) True or false, with a reason: if t is a variable, then t is free for x in every formula.

(d) True or false, with a reason: $(\forall x\,P(x) \wedge Q(x))[c/x]$ is $\forall x\,P(x) \wedge Q(c)$.

Solution. (a) (i) The single occurrence of x, in $P(x)$, is bound by $\forall x$ (the x after $\forall$ belongs to the quantifier node and is not an occurrence); the occurrence of y in $Q(y)$ is free (no quantifier on y). Free variables: y. Not a sentence.

(ii) The occurrence of x is bound by $\exists x$, that of y by $\forall y$. No free variables: a sentence.

(iii) By shorthand rule 2 this is $(\forall x\,(x < y) \wedge \exists y\,(y < x))$. In the left part, x is bound by $\forall x$ and y is free. In the right part, y is bound by $\exists y$ and x is free (the $\forall x$ governs only the left part). Free variables: x and y, each with one free and one bound occurrence. Not a sentence.

(iv) Every occurrence of ε lies below $\forall \varepsilon$, and both occurrences of δ lie below $\exists \delta$; 0 is a constant. A sentence.

(b) (i) A term, by (Tm2) twice. (ii) A formula, atomic by (Fm1): the predicate R applied to the terms $f(x)$ and c. (iii) Neither: $R(x, c)$ is a formula, not a term, and (Tm2) applies f only to a term. (iv) Neither: (Fm4) puts a **variable** after $\forall$, and c is a constant symbol.

(c) False. The variable y is not free for x in $\exists y\,(y \neq x)$, since the free occurrence of x lies in the scope of $\exists y$. (A variable t **is** free for x when t is x itself, or when t does not appear in the formula.)

(d) True. The formula is $(\forall x\,P(x) \wedge Q(x))$ by shorthand rule 2. The $\wedge$ passes the substitution to both parts; $\forall x\,P(x)$ is unchanged because its quantifier is on x, and $Q(x)$ becomes $Q(c)$.

Exercise. A2: Drill

Answer each part in a line or two.

(a) In $\forall y\,(x < y) \wedge \exists x\,(x = y)$, which occurrences are free, and what are the free variables?

(b) Is $\forall x\,\exists y\,(x + y = 0)$ a sentence?

(c) Someone writes $\forall x\,P(x) \to Q(x)$ for "every P is a Q". Which occurrence of x is free, and what is the repair?

(d) Repair $\forall \varepsilon > 0\ \exists \delta > 0\ (|x - a| < \delta \to |f(x) - f(a)| < \varepsilon)$, meant as "$f$ is continuous at a", with a a variable here (not a constant symbol), so that a is its only free variable.

(e) Compute $(x \leq y \wedge \exists y\,(y < x))[3/x]$.

(f) Compute $(\forall x\,(x \geq y))[0/y]$.

(g) Compute $(\forall x\,P(x) \to P(x))[c/x]$.

(h) Compute $((x + y = y + x)[y/x])[x/y]$. Is the result the formula with x and y swapped?

(i) Compute $(P(x, y)[f(y)/x])[c/y]$.

(j) Is x free for y in $\exists x\,(x > y)$?

(k) Is $z + 1$ free for y in $\exists x\,(x > y)$? Compute.

(l) Put x in for y in $\exists x\,(x > y)$, renaming first.

Solution.

(a) The x of $x < y$ and the y of $x = y$ are free; the y of $x < y$ is bound by $\forall y$ and the x of $x = y$ by $\exists x$. Free variables: x, y.

(b) Yes: every occurrence is bound.

(c) By shorthand rule 2 the $\forall x$ governs only $P(x)$, so the formula is $(\forall x\,P(x)) \to Q(x)$ and the x of $Q(x)$ is free. Repair: $\forall x\,(P(x) \to Q(x))$.

(d) As written, x is free, so the formula speaks of one fixed x. Repair: $\forall \varepsilon > 0\ \exists \delta > 0\ \forall x\,(|x - a| < \delta \to |f(x) - f(a)| < \varepsilon)$; f is a function symbol, not a variable.

(e) $3 \leq y \wedge \exists y\,(y < 3)$.

(f) $\forall x\,(x \geq 0)$.

(g) $\forall x\,P(x) \to P(c)$ (shorthand rule 2: the $\forall x$ governs only the first $P(x)$).

(h) $y + y = y + y$, then $x + x = x + x$. No: the swap is $y + x = x + y$.

(i) $P(f(y), y)$, then $P(f(c), c)$: the second substitution also reaches the y inside the term put in by the first.

(j) No: the free y lies in the scope of $\exists x$, and x is the variable of the term.
(k) Yes: $\exists x\,(x > z + 1)$.
(l) Rename x to z: $\exists z\,(z > y)$; then $\exists z\,(z > x)$.

5.2.8.2 B. Practice

Exercise. B1: Computing substitutions

Compute each substitution. Where the term is not free for the variable, say which quantifier would capture it, rename first, and then substitute.

(a) $(P(x) \wedge R(x, y))[f(y)/x]$.
(b) $(\forall x\, R(x, y) \to P(x))[c/x]$.
(c) $(\forall y\,(x < y))[(y + 1)/x]$.
(d) $(\exists x\,(x = 2 \cdot y))[x/y]$.
(e) $(\exists y\,(y > x) \wedge Q(y))[(x + 1)/y]$.

Solution. *(a)* No quantifiers, so $f(y)$ is free for x. Every occurrence of x is replaced: $P(f(y)) \wedge R(f(y), y)$. The y inside $f(y)$ and the y already in $R(x, y)$ now stand side by side; with no quantifier around, both are the same free y, as intended.

(b) By shorthand rule 2 the formula is $(\forall x\, R(x, y) \to P(x))$. The constant c has no variables, so it is free for x. The left part is unchanged (quantifier on x); the free x in $P(x)$ becomes c: $\forall x\, R(x, y) \to P(c)$.

(c) Not free for: the free occurrence of x lies in the scope of $\forall y$, and y is a variable of $y + 1$. Rename y to z, which appears nowhere: $\forall z\,(x < z)$. Then substitute: $\forall z\,(y + 1 < z)$. (Without renaming, the definition would give $\forall y\,(y+1 < y)$, which says nothing about $y+1$.)

(d) Not free for: the free occurrence of y lies in the scope of $\exists x$, and the term is x. Rename x to z: $\exists z\,(z = 2 \cdot y)$, "$y$ doubled is something", then substitute: $\exists z\,(z = 2 \cdot x)$. (Without renaming, the result would be $\exists x\,(x = 2 \cdot x)$, "some number is its own double", a claim about no particular number.)

(e) By shorthand rule 2 the formula is $(\exists y\,(y > x) \wedge Q(y))$. The occurrence of y in $y > x$ is bound, and the one in $Q(y)$ is free and not in the scope of any quantifier; so $x + 1$ is free for y (the $\exists y$ above the bound y is irrelevant, since that occurrence is not replaced). The result is $\exists y\,(y > x) \wedge Q(x + 1)$.

Exercise. B2: The set of free variables

For a term t, let $\mathrm{Var}(t)$ be the set of variables of t, and for a formula φ, let $\mathrm{FV}(\varphi)$ be the set of its free variables. (Here $\cup$ is the union of two sets, $\setminus$ removes elements, and $\emptyset$ is the empty set, all in the everyday sense.)

(a) Write down clauses that define $\mathrm{Var}(t)$ by recursion on terms and $\mathrm{FV}(\varphi)$ by recursion on formulas, one clause for each case of Definition 5.2.1 and of Definition 5.2.2.
(b) Use your clauses to compute FV of $\forall x\,(x < y) \wedge \exists y\,(y < x)$ and of $\forall x\,\exists y\,(x + y = z)$,

showing each step.

Solution. *(a)* For terms:

$$\mathrm{Var}(x) = \{x\} \text{ for a variable } x, \qquad \mathrm{Var}(c) = \emptyset,$$
$$\mathrm{Var}\big(f(t_1, \ldots, t_n)\big) = \mathrm{Var}(t_1) \cup \cdots \cup \mathrm{Var}(t_n).$$

For formulas:

$$\mathrm{FV}\big(R(t_1, \ldots, t_n)\big) = \mathrm{Var}(t_1) \cup \cdots \cup \mathrm{Var}(t_n),$$
$$\mathrm{FV}(t_1 = t_2) = \mathrm{Var}(t_1) \cup \mathrm{Var}(t_2), \qquad \mathrm{FV}(\bot) = \emptyset,$$
$$\mathrm{FV}(\neg\varphi) = \mathrm{FV}(\varphi), \qquad \mathrm{FV}\big((\varphi \circ \psi)\big) = \mathrm{FV}(\varphi) \cup \mathrm{FV}(\psi),$$
$$\mathrm{FV}(\forall x\, \varphi) = \mathrm{FV}(\exists x\, \varphi) = \mathrm{FV}(\varphi) \setminus \{x\}.$$

These are the four lines of the recursive description of "free" (in the fold after the examples of free and bound occurrences), rewritten as sets: $x \in \mathrm{FV}(\varphi)$ exactly when x is free in φ. By structural recursion (carried over to terms and to first-order formulas), each set of clauses defines exactly one function.

(b) First formula. $\mathrm{FV}(x < y) = \{x, y\}$, so $\mathrm{FV}(\forall x\,(x < y)) = \{x, y\} \setminus \{x\} = \{y\}$. $\mathrm{FV}(y < x) = \{x, y\}$, so $\mathrm{FV}(\exists y\,(y < x)) = \{x\}$. The conjunction gives $\{y\} \cup \{x\} = \{x, y\}$, which agrees with A1(a)(iii).

Second formula. $\mathrm{FV}(x + y = z) = \mathrm{Var}(x + y) \cup \mathrm{Var}(z) = \{x, y, z\}$. Then $\mathrm{FV}(\exists y\,(x + y = z)) = \{x, z\}$, and $\mathrm{FV}(\forall x\, \exists y\,(x + y = z)) = \{z\}$. So the formula is an open sentence in z alone (over $\mathbb{R}$ it is true for every z: given x, take $y = z - x$).

Exercise. B3: Swapping two letters

Let $\varphi = \exists z\,(R(x, z) \wedge R(z, y))$.

(a) Compute $(\varphi[y/x])[x/y]$, and say whether it is φ with x and y swapped.

(b) Swap x and y in φ with three substitutions through a letter u that appears nowhere, checking at each step that the term is free for the variable.

(c) Why would z not do as the go-between?

Solution. *(a)* $\varphi[y/x] = \exists z\,(R(y, z) \wedge R(z, y))$, and then $[x/y]$ gives $\exists z\,(R(x, z) \wedge R(z, x))$. The swap is $\exists z\,(R(y, z) \wedge R(z, x))$, a different formula, and it says something different: over $\mathbb{N}$ with R read as $<$ and x, y naming $2, 0$, the swap is true ($z = 1$), while the computed formula, "some z is above 2 and below 2", is false.

(b) $\varphi[u/x] = \exists z\,(R(u, z) \wedge R(z, y))$; then $[x/y]$ gives $\exists z\,(R(u, z) \wedge R(z, x))$; then $[y/u]$ gives $\exists z\,(R(y, z) \wedge R(z, x))$, the swap. At each step the only quantifier above the replaced occurrence is $\exists z$, and the term $(u,\ x,\ y)$ is not z, so it is free.

(c) z is not free for x in φ: the x of $R(x, z)$ lies inside $\exists z$, so $\varphi[z/x]$ would be captured and say $\exists z\,(R(z, z) \wedge R(z, y))$. The go-between must appear nowhere in φ, bound letters included ("appear", in "Free and bound: drawing the arrows").

Worked examples of these types: "One substitution at a time"; the practical rule "Before

you substitute".

Marking scheme. (10 points)

- 3: (a), both steps computed and the swap written, with the difference shown (by the strings, or by one reading in which they differ).
- 4: (b), the three substitutions in the right order, each result correct, and the free-for check stated once for each step.
- 3: (c), capture named, with the captured formula written.

Fatal error (0 for the part): treating $[y/x]$ and $[x/y]$ as happening at the same time.

5.2.8.3 C. Going deeper

Exercise. C1: Challenge: What substitution does to the free variables

Let φ be a formula, x a variable and t a term, with t free for x in φ, and with x free in φ.

(a) Prove that the free variables of $\varphi[t/x]$ are the free variables of φ other than x, together with the variables of t:

$$\mathrm{FV}(\varphi[t/x]) = (\mathrm{FV}(\varphi) \setminus \{x\}) \cup \mathrm{Var}(t),$$

with the notation of B2. Prove it in the following form, by structural induction on φ: for every φ in which t is free for x, *if* $x \in \mathrm{FV}(\varphi)$, *then* $\mathrm{FV}(\varphi[t/x]) = (\mathrm{FV}(\varphi)\setminus\{x\})\cup\mathrm{Var}(t)$; *if* $x \notin \mathrm{FV}(\varphi)$, *then* $\varphi[t/x] = \varphi$. The term version, the atomic case and the connective cases are done below; prove the case of $\forall y\, \varphi$ (the case of $\exists y\, \varphi$ is the same).

(b) Show, with the capture example $\exists y\, (y > x)$, $t = y + 1$, that the conclusion can fail when t is not free for x. Show also, with an example of your own, that it can fail when x is not free in φ.

Given for (a). *Terms.* Let s be a term in which x occurs. Then $\mathrm{Var}(s[t/x]) = (\mathrm{Var}(s) \setminus \{x\}) \cup \mathrm{Var}(t)$. By induction on s: if s is the variable x, then $s[t/x] = t$ and both sides are $\mathrm{Var}(t)$. (A constant, or a variable other than x, does not contain x.) If $s = f(s_1, \ldots, s_n)$, then $s[t/x] = f(s_1[t/x], \ldots, s_n[t/x])$; for each s_i containing x, the induction hypothesis gives $\mathrm{Var}(s_i[t/x]) = (\mathrm{Var}(s_i) \setminus \{x\}) \cup \mathrm{Var}(t)$, and for each s_i not containing x, $s_i[t/x] = s_i$ (as in the proof of the claim in the text). At least one s_i contains x, so the union over i is $(\mathrm{Var}(s) \setminus \{x\}) \cup \mathrm{Var}(t)$.

Formulas. We prove, by structural induction on φ, the statement $S(\varphi)$: *if t is free for x in φ, then* $\mathrm{FV}(\varphi[t/x]) = (\mathrm{FV}(\varphi)\setminus\{x\})\cup\mathrm{Var}(t)$ *when* $x \in \mathrm{FV}(\varphi)$, *and* $\varphi[t/x] = \varphi$ *when* $x \notin \mathrm{FV}(\varphi)$. The second half is the claim proved in the text, for every φ; so in each case below we assume $x \in \mathrm{FV}(\varphi)$ and prove the equation. Write $E(\chi)$ for the set $(\mathrm{FV}(\chi) \setminus \{x\}) \cup \mathrm{Var}(t)$.

Atomic. Let $\varphi = R(s_1, \ldots, s_n)$ with x occurring in some s_i. By the clause for atomic formulas, $\mathrm{FV}(\varphi[t/x])$ is the union over i of the sets $\mathrm{Var}(s_i[t/x])$. Exactly as in the term case for $f(s_1, \ldots, s_n)$ (the result for terms for each s_i containing x, $s_i[t/x] = s_i$ for the others, and at least one s_i containing x), this union is $E(\varphi)$. The same for $s_1 = s_2$. ($\bot$ has no free variables.)

Negation. $\mathrm{FV}((\neg\varphi)[t/x]) = \mathrm{FV}(\varphi[t/x]) = E(\varphi) = E(\neg\varphi)$, by the induction hypothesis (t is free for x in φ, and x is free in φ).

Binary connective. Let x be free in $(\varphi \circ \psi)$; then t is free for x in both parts, and x is free in

at least one. If it is free in both, the induction hypothesis gives $E(\varphi) \cup E(\psi) = E((\varphi \circ \psi))$. If it is free in φ only, then $\psi[t/x] = \psi$, and $E(\varphi) \cup \mathrm{FV}(\psi) = E((\varphi \circ \psi))$, because $x \notin \mathrm{FV}(\psi)$. Symmetrically if it is free in ψ only.

Solution (Hint 1). When x is free in $\forall y\, \varphi$, what does "free for" tell you about y and t?

Solution (Hint 2). Start from $\mathrm{FV}((\forall y\, \varphi)[t/x]) = \mathrm{FV}(\varphi[t/x]) \setminus \{y\}$ and apply the induction hypothesis to φ. Then see where removing y from the resulting set needs something about y and $\mathrm{Var}(t)$.

Solution. (a) The term version and the atomic, negation and connective cases are given in the statement. It remains to treat the quantifiers.

Quantifier. If x is not free in $\forall y\, \varphi$, then $(\forall y\, \varphi)[t/x] = \forall y\, \varphi$ by the claim of the text. Now let x be free in $\forall y\, \varphi$. Then y is not x and x is free in φ; and since t is free for x in $\forall y\, \varphi$ while x is free there, $y \notin \mathrm{Var}(t)$ and t is free for x in φ (the recursive description of "free for"). So

$$\begin{aligned} \mathrm{FV}((\forall y\, \varphi)[t/x]) &= \mathrm{FV}(\varphi[t/x]) \setminus \{y\} = ((\mathrm{FV}(\varphi) \setminus \{x\}) \cup \mathrm{Var}(t)) \setminus \{y\} \\ &= (\mathrm{FV}(\varphi) \setminus \{x, y\}) \cup \mathrm{Var}(t) = E(\forall y\, \varphi), \end{aligned}$$

where the third equality uses $y \notin \mathrm{Var}(t)$, and the last uses $\mathrm{FV}(\forall y\, \varphi) = \mathrm{FV}(\varphi) \setminus \{y\}$. The case of $\exists y$ is the same.

This proves $S(\varphi)$ for every formula, and (a) is its first half.

(b) With $\varphi = \exists y\, (y > x)$ and $t = y + 1$: $\mathrm{FV}(\varphi) = \{x\}$, so the equation would give $\mathrm{FV}(\varphi[t/x]) = \{y\}$. But $\varphi[t/x] = \exists y\, (y > y + 1)$, whose set of free variables is $\emptyset$. The y of the term was captured, so it is not free in the result; the quantifier case of the proof fails because $y \in \mathrm{Var}(t)$.

For the second hypothesis, take $\varphi = P(z)$ and $t = y$. Then x is not free in φ, and t is free for x (there are no quantifiers). Now $\varphi[t/x] = P(z)$, with free variables $\{z\}$, while the equation would give $\{z\} \cup \{y\} = \{y, z\}$. Without a free x, nothing is put in, so the variables of t do not appear.

Marking scheme. (10 points)

- 3: the quantifier case reduced to $\mathrm{FV}(\varphi[t/x]) \setminus \{y\}$ by the definitions.
- 3: the use of "free for" to get $y \notin \mathrm{Var}(t)$ and t free for x in φ, naming where it is used.
- 1: the case "x not free in $\forall y\, \varphi$" handled by the claim of the text.
- 3: both counterexamples of (b), each checked by computing both sides.

Fatal errors (0 for the step): using the equation $\mathrm{FV}(\chi[t/x]) = (\mathrm{FV}(\chi) \setminus \{x\}) \cup \mathrm{Var}(t)$ for a part χ in which x is not free (it fails there; see the second example in (b)); removing y from $\mathrm{Var}(t)$ without justification.

Exercise. C2: Renaming is not swapping

(a) Show that $\forall x\,\exists y\,(x < y)$ and $\forall y\,\exists x\,(y < x)$ are the same sentence up to renaming bound variables: give a sequence of renamings, each allowed by the condition of the text, leading from the first to the second.

(b) Show that $\forall x\,\exists y\,(x < y)$ and $\exists y\,\forall x\,(x < y)$ are **not** the same up to renaming. *Hint: what does a renaming leave unchanged?*

Solution. (a) A renaming needs a variable that does not appear in the formula, so the two letters cannot be exchanged in one step; go through a third letter z.

1. In $\forall x\,\exists y\,(x < y)$, rename the bound x to z, which does not appear: $\forall z\,\exists y\,(z < y)$.
2. Rename the bound y to x, which no longer appears: $\forall z\,\exists x\,(z < x)$.
3. Rename the bound z to y, which no longer appears: $\forall y\,\exists x\,(y < x)$.

Each step renames one quantifier and the occurrences it binds, to a letter absent from the formula at that moment. The result is the second sentence. (Over $\mathbb{N}$ both say "there is no largest natural number", in two spellings.)

(b) A renaming changes only which **letters** are used. It does not change the parse tree's shape, which quantifier symbol ($\forall$ or $\exists$) sits at which node, or which quantifier node binds which occurrence. In $\forall x\,\exists y\,(x < y)$ the root is a $\forall$ node; in $\exists y\,\forall x\,(x < y)$ the root is an $\exists$ node. No sequence of renamings turns a $\forall$ node into an $\exists$ node, so no sequence of renamings leads from one sentence to the other.

The two sentences also say different things: over $\mathbb{N}$, the first is true (take $y = x + 1$) and the second false (for any y, take $x = y$). This is Example (Swapping two quantifiers) again. Renaming changes the **names** of the placeholders; swapping changes the **order** in which they are chosen, and the order is what carries the meaning.

5.3 Using "For Every": Instantiation

Section 5.2 made the language exact and defined the operation $\varphi[t/x]$, "put the term t in for the variable x", with its one hazard, capture. We can now give the first quantifier rule. It is the rule for **using** a universal statement: whatever holds for every object holds for the object named by any term we care to write. The rule is called $\forall$E. It is the rule behind every "by the theorem, applied to $x = 2$". With it alone, universal statements can be used but not proved, and the section meets exactly that gap ("A stall" below), which Section 5.4 fills.

5.3.1 Applying a statement about everything

Every application of a general fact to a particular case is one step of logic, and so far that step has been taken on trust. Three earlier places took it.

- Section 1.3 used the theorem "if an integer is divisible by 2 and by 3, then it is divisible by 6" at $n = 42$, taking the step from "every integer" to the one integer 42 on trust.
- In Example (A number below every positive number) (Section 2.5) the hypothesis was "$x \le \varepsilon$ for every $\varepsilon > 0$", and the proof applied it to the one number $\varepsilon = x/2$.
- Section 5.1 opened with the argument "every positive real has a square root; 2 is positive; so 2 has a square root", which a truth table calls invalid because it sees only three unrelated atoms.

The same step happens in each: a statement about **every** object is applied to **one** object, named by a term. In the language of Section 5.2, the first premise of the square-root argument is $\forall x\,(P(x) \to S(x))$, and applying it to 2 gives

$$(P(x) \to S(x))[2/x] \;=\; P(2) \to S(2),$$

which is part (a) of Example (Four substitutions). From there $\to$E finishes. All that is missing is a rule that licenses the first step.

5.3.2 The rule

A universal statement may be applied to any term: remove the quantifier and put the term in for the variable.

Definition 5.3.1. Universal Elimination

i	$\forall x\,\varphi$	
k	$\varphi[t/x]$	$\forall$E i

Here φ is a formula, x a variable, and t **any** term, subject to one side condition: t **is free for** x **in** φ (Definition 5.2.6). Line i must be available at line k (the scope rule of Definition 1.4.2). This rule is called $\forall$E (**universal elimination**, or **instantiation**), and $\varphi[t/x]$ is an **instance** of $\forall x\,\varphi$.

In words: if a line says that φ holds for every x, then any later line, where that line is still available, may write what φ says about t. The rule cites one line and no box.

- **Schema:** the block above: one cited line, no box.
- **Reading:** drop the $\forall x$, put one term in for every free x, and check that no letter of the term gets caught by a quantifier inside.
- **Example:** from $\forall x\,(P(x) \to S(x))$, "$\forall$E 1" at the term 2 writes $P(2) \to S(2)$.

Read the definition clause by clause; exercises cite the clauses by number.

1. **The cited line's main symbol is** $\forall x$, with all of φ as its scope: $\forall x\,P(x) \to Q(c)$ is an implication; a $\forall$ buried inside another connective must first be taken out by that connective's rule, when it can be ($\wedge$E takes $\forall x\,P(x)$ out of $\forall x\,P(x) \wedge Q(c)$), and a $\forall$ in the hypothesis of an implication cannot be instantiated at all.
2. **Any term:** a constant, a variable (x itself included) or a compound term. Nothing about it needs to be established first, since every term names an object of the domain.
3. **Every free occurrence, by the same term:** the new line is $\varphi[t/x]$; we do not choose which occurrences change.
4. **Free for** x (Definition 5.2.6): no letter of t is caught by a quantifier of φ. This is the rule's only restriction.
5. **Line** i **is available** (the scope rule of Definition 1.4.2), as for every rule.

When the universal statement has a bound, $\forall x \in A\;\varphi$, using it at t also needs the line $t \in A$.

Quick check. Line 1 is $\forall x\,(x + 0 = x)$. Which of $3 + 0 = 3$, $y + 0 = y$, $3 + 0 = x$ may line 2 be, by $\forall$E?

Solution. The first two, at the terms 3 and y; y is free for x, since the formula has no quantifier. The third is not an instance: $\varphi[t/x]$ puts the **same** term in for every free x, and $3 + 0 = x$ replaced one x and left the other.

↻ **Second reading** · *A quantifier that binds nothing*

If x is not free in φ, then $\varphi[t/x] = \varphi$ for every term t (the claim proved in Section 5.2), and ∀E simply removes a quantifier that bound nothing: from $\forall x\, P(c)$ it writes $P(c)$. Legal, and dull.

One line, many uses. The cited line stays where it is, available as before. A derivation may instantiate the same universal line as often as it likes, at different terms or the same one. Worked derivation (d) below uses one premise twice.

Why this rule. It says no more than the word "every". If every object in the domain satisfies φ, then so does the object that t names, whichever object that is. A term always names something, since the domain is never empty (Section 5.1), so the instance always has something to be about.

Why "free for". The instance must say about t what φ says about x. Section 5.2 showed that when a variable of t is captured, $\varphi[t/x]$ can say something quite different. The condition rules capture out, and the subsection on common errors below shows the rule going wrong without it.

Quick check. Line 1 of a derivation is $\forall x\, \exists z\, (z > x + y)$. For which terms t may ∀E write $(\exists z\, (z > x + y))[t/x]$ on line 2? Is $t = y$ allowed? Is $t = z + 1$?

Solution. The only free occurrence of x in $\exists z\, (z > x + y)$ lies in the scope of $\exists z$, and that is the only quantifier above it. So t is free for x exactly when z is not a variable of t: **any term not containing** z is allowed. The term y is allowed, and gives $\exists z\, (z > y + y)$; the variable y already occurs free in the line, and that does no harm. The term $z + 1$ is not allowed: it would give $\exists z\, (z > z + 1 + y)$, where the z of the term has been captured. Nor is z itself. To instantiate at z, rename the bound z first (Section 5.2), say to w: from $\forall x\, \exists w\, (w > x + y)$, ∀E at z gives $\exists w\, (w > z + y)$.

A long "and", taken apart. In world W of Section 5.1, $\forall x\, \varphi$ says $\varphi[a/x] \wedge \varphi[b/x] \wedge \varphi[c/x] \wedge \varphi[d/x]$, and ∀E at b does what ∧E does: it takes out one conjunct. Over $\mathbb{N}$ the conjunction cannot be written; ∀E takes out any conjunct, at any term, without writing the rest.

5.3.3 Four derivations

The worked derivations below use ∀E and the rules of Chapter 1. Each has an Idea that says why each rule was chosen, continuing the search method of Section 3.1: a goal-directed step looks at the goal's main symbol, a premise-directed step at a resource's. The new question is the one that ∀E raises every time: **at which term?**

Example. Instantiation at work

Give a derivation of each sequent.

(a) $\forall x\,(P(x) \to Q(x)),\ P(c) \vdash Q(c)$.

(b) $\forall x\,\forall y\,R(x,y) \vdash R(c,c)$.

(c) $\forall x\,\forall y\,(R(x,y) \to R(y,x)),\ R(a,b) \vdash R(b,a)$.

(d) $\forall x\,(P(x) \to P(f(x))),\ P(c) \vdash P(f(f(c)))$.

Solution.

(a) *Idea.* The goal $Q(c)$ has no connective or quantifier, so no introduction rule produces it, and the search turns to the premises. The premise $P(c)$ cannot be taken apart. The premise $\forall x\,(\dots)$ begins with $\forall$, so it can be used only by $\forall$E. At which term? The goal mentions the term c, so instantiate at c. That gives an implication whose hypothesis is a premise and whose conclusion is the goal.

1	$\forall x\,(P(x) \to Q(x))$	
2	$P(c)$	
3	$P(c) \to Q(c)$	$\forall$E 1
4	$Q(c)$	$\to$E 3, 2

Line 3 is $(P(x) \to Q(x))[c/x]$; the constant c has no variables, so it is free for x. Read c as 2, $P(x)$ as "x is positive", and $Q(x)$ as "x has a square root" (Section 5.1 wrote S for this predicate), and the derivation is the square-root argument of Section 5.1, in the four lines promised there. The argument that truth tables rejected is derivable once the rules can see inside its sentences.

(b) *Idea.* The premise has two quantifiers, and $\forall$E removes one at a time, the outermost first: the main symbol of the premise is $\forall x$, and only after it is gone is $\forall y$ the main symbol. The goal $R(c,c)$ shows the term for both.

1	$\forall x\,\forall y\,R(x,y)$	
2	$\forall y\,R(c,y)$	$\forall$E 1
3	$R(c,c)$	$\forall$E 2

Line 2 is $(\forall y\,R(x,y))[c/x]$: the quantifier $\forall y$ is on a variable other than x, so the substitution passes through it (Definition 5.2.5). Line 3 instantiates y at c as well. Nothing forbids putting the same term in for two different variables; "for all x and y" includes the case $x = y$.

(c) *Idea.* Work backwards from the goal $R(b,a)$. The only line that can produce it is the first premise, through an instance of its inside, $R(x,y) \to R(y,x)$, followed by $\to$E. Which instance has conclusion $R(b,a)$? Matching $R(y,x)$ with $R(b,a)$ gives $y \coloneqq b$ and $x \coloneqq a$. Then the hypothesis of that instance is $R(a,b)$, which is the second premise. So the terms are read off the goal: a for x first (the outer quantifier), then b for y.

1	$\forall x\,\forall y\,(R(x,y) \to R(y,x))$	
2	$R(a,b)$	
3	$\forall y\,(R(a,y) \to R(y,a))$	$\forall$E 1
4	$R(a,b) \to R(b,a)$	$\forall$E 3
5	$R(b,a)$	$\to$E 4, 2

A reader who matches the goal $R(b,a)$ against the hypothesis $R(x,y)$ instead of the conclusion $R(y,x)$ puts b for x and a for y, and gets $R(b,a) \to R(a,b)$. That is a legal line, but its hypothesis is the goal itself, so $\to$E cannot fire. Match the goal against the conclusion of the implication, since that is the part $\to$E delivers.

(d) *Idea, forwards.* The premise $P(c)$ is the hypothesis of the instance of the first premise at c, so instantiating at c and firing gives $P(f(c))$. Now $P(f(c))$ is the hypothesis of the instance at $f(c)$, so the same premise, instantiated again, fires again and gives $P(f(f(c)))$, the goal.

Idea, backwards. The goal $P(f(f(c)))$ has the shape $P(f(x))$, the conclusion of the inside of the first premise, with $x := f(c)$. So the last $\forall$E must be at the **compound** term $f(c)$, and its hypothesis $P(f(c))$ becomes the new goal. That has the shape $P(f(x))$ with $x := c$, and its hypothesis is the premise $P(c)$.

1	$\forall x\,(P(x) \to P(f(x)))$	
2	$P(c)$	
3	$P(c) \to P(f(c))$	$\forall$E 1
4	$P(f(c))$	$\to$E 3, 2
5	$P(f(c)) \to P(f(f(c)))$	$\forall$E 1
6	$P(f(f(c)))$	$\to$E 5, 4

Line 5 cites line 1 again: a universal line can be instantiated as often as needed. Its term $f(c)$ contains no variable, so it is free for x, and $(P(x) \to P(f(x)))[f(c)/x]$ is $P(f(c)) \to P(f(f(c)))$: **every** x is replaced by the whole term $f(c)$, including the one inside $f(x)$.

An arithmetic reading: over $\mathbb{N}$, let $f(n)$ be $n+2$, let $P(n)$ say "n is even", and let c be 0. The first premise says that adding 2 to an even number gives an even number, the second that 0 is even, and the derivation concludes that $f(f(0)) = 4$ is even. Each further step up the even numbers would cost two more lines. Reaching "every n of the form $f(\cdots f(0)\cdots)$" at once is a job for induction, not for this rule.

In each derivation, the Idea chose the term of every $\forall$E by matching: the goal, or a hypothesis still needed, has a definite shape, and the instance is chosen to have that shape. Instantiating at a term that nothing asks for is legal, and it fills a derivation with lines that lead nowhere.

5.3.4 Checking a line of instantiation

Choosing the term can take thought; checking a line never does. For a line justified "$\forall$E i", ask four questions.

1. Is line i available here (Definition 1.4.2)?

2. Is the main symbol of line i a universal quantifier? If so, read off its variable x and its scope φ.
3. Is this line $\varphi[t/x]$ for some term t? Find t by comparing the two lines at a free occurrence of x in φ; then check that **every** free occurrence of x became that same t, and that nothing else changed.
4. Is t free for x in φ?

Run on line 5 of (d): line 1 is available, since no box contains it; its main symbol is $\forall x$, with scope $P(x) \to P(f(x))$; comparing gives $t = f(c)$, and both occurrences of x became $f(c)$; and $f(c)$ has no variables, so it is free for x. The line passes.

5.3.5 How this rule looks in a mathematical proof

Let a be a real number. A proof might say:

> Since $y^2 \geq 0$ for every real y, applying this to $a - 1$ gives $(a-1)^2 \geq 0$, that is, $a^2 - 2a + 1 \geq 0$.

The first half of the sentence is one ∀E, at the compound term $a - 1$. The second half rewrites $(a-1)^2$ by school algebra. In a derivation such a step is written as one line justified "arith", for **arithmetic**:

1	$\forall y\,(y^2 \geq 0)$	
2	$(a-1)^2 \geq 0$	∀E 1
3	$a^2 - 2 \cdot a + 1 \geq 0$	arith 2

Line	Sentence
1	"Since $y^2 \geq 0$ for every real y, …" (the general fact)
2	"… applying this to $a - 1$ gives $(a-1)^2 \geq 0$, …" (∀E at $a - 1$)
3	"… that is, $a^2 - 2a + 1 \geq 0$." (arith)

A line justified "arith" records a fact of school arithmetic or algebra, used on the lines it cites; here the fact is that $(u-1)^2$ and $u^2 - 2u + 1$ are the same number for every real u. It is **not a new rule**, and the facts of school algebra are used as facts.

↻ **Second reading** · *What an arith line stands for*

Written out in full, an arith line is a universally quantified arithmetic statement, here

$$\forall u\,\big((u-1)^2 \geq 0 \to u^2 - 2 \cdot u + 1 \geq 0\big),$$

taken as one more premise, instantiated by ∀E at the term on the line (here a) and applied by →E to the cited line. That is exactly the pattern of (a) above. The book writes it as one line because spelling it out every time would bury the logic under algebra; its bound variables are letters that appear nowhere else in the derivation, and that keeps every term free for them.

Section 5.4 will turn this three-line derivation into a proof of $x^2 - 2x + 1 \geq 0$ for **every** real x. That needs a rule this section does not have.

5.3.6 Common errors

Three mistakes are worth seeing written out, each with a reading in which the premise is true and the offending line false.

Capture. The premise $\forall x\, \exists y\, (y \neq x)$ says, over $\mathbb{N}$, that for every number there is a different one; it is true, since $x + 1 \neq x$. Instantiating "at y" without looking writes this:

1	$\forall x\, \exists y\, (y)$	$\neg$E **??**
2	$\exists y\, (y)$	$\neg$E **??**

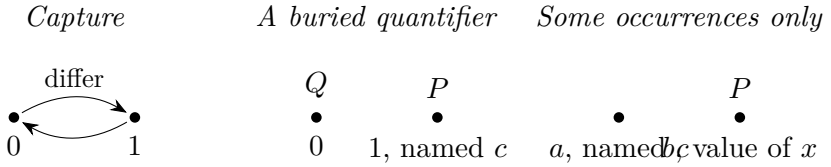
Line 2 says that some number differs from itself, which is false. The step breaks the side condition: the free occurrence of x in $\exists y\, (y \neq x)$ lies in the scope of $\exists y$, and y is a variable of the term y, so y is not free for x. This is the capture example of Section 5.2, now inside a derivation. The repair is to rename first: $\forall x\, \exists z\, (z \neq x)$ is the same premise up to renaming, and $\forall$E at y gives $\exists z\, (z \neq y)$, "some number differs from y", which is what was meant.

Warning. $\forall$E at a term that is not free for x can turn a true premise into a false line, and the definition of $\varphi[t/x]$ gives no signal: it returns a formula either way. Over $\mathbb{N}$, from the true $\forall x\, \exists y\, (y \neq x)$, the "instance at y" is $\exists y\, (y \neq y)$, which is false. Before instantiating at a term with variables, look for quantifiers on those variables inside φ.

A buried quantifier. From the line $\forall x\, P(x) \to Q(c)$ one may **not** write $P(c) \to Q(c)$. The main symbol of the line is $\to$, and the $\forall x$ governs only the hypothesis $\forall x\, P(x)$. The line says "if **every** x has P, then $Q(c)$"; the false step reads it as "if c has P, then $Q(c)$", which is far stronger. Over $\mathbb{N}$, read $P(x)$ as $x \geq 1$, $Q(x)$ as $x = 0$, and c as 1. The line is true, because its hypothesis "every natural number is at least 1" is false (0 is not). The false step's result, "if $1 \geq 1$, then $1 = 0$", is false.

Replacing only some occurrences. Over $\mathbb{R}$, $\forall x\, (x \leq x^2 \vee x < 1)$ is true: a real $x \geq 1$ has $x^2 = x \cdot x \geq x$, and every other real has $x < 1$. Writing $2 \leq x^2 \vee x < 1$ from it replaces the first x and leaves the others. That is not $\varphi[2/x]$, which is $2 \leq 2^2 \vee 2 < 1$: the substitution replaces every free occurrence. The half-replaced line still has x free, and when x names 1 it reads $2 \leq 1 \vee 1 < 1$, which is false.

Each error has a reading, a small one, in which the premise is true and the faulty line false.



Three tiny worlds. In each, the premise is true and the faulty line false. Capture: $\forall x\, \exists y\, (y \neq x)$ is true, since each dot has the other; line 2, $\exists y\, (y \neq y)$, is false, since no dot differs from itself. A buried quantifier: $\forall x\, P(x) \to Q(c)$ is true, since $\forall x\, P(x)$ fails at 0; $P(c) \to Q(c)$ is false, since 1 has P and not Q. Some occurrences only: $\forall x\, (P(x) \vee \neg P(x))$ is true, since at each dot one side holds; the half-replaced $P(c) \vee \neg P(x)$, which put c in for the first x only, is false, since

a lacks P and x names b, which has it. The instance $P(c) \vee \neg P(c)$ is true, as every instance must be.

Quick check. Which numbered clause of ∀E does each faulty step of the figure break?

Solution. Capture breaks clause 4 (y is not free for x); the buried quantifier breaks clause 1 (the main symbol is $\rightarrow$); replacing some occurrences only breaks clause 3 (the line is not $\varphi[c/x]$).

5.3.7 A stall

Try the sequent

$$\forall x\,(P(x) \wedge Q(x)) \vdash \forall x\,P(x).$$

It is plainly correct: if every x has both properties, then every x has the first. Search for a derivation. The goal begins with ∀, and no rule so far is made for proving a ∀. Apart from ⊥E (Definition 2.1.1) and RAA (Definition 2.5.1), which reach a line only by way of a contradiction, a line beginning with ∀ can only be taken out of a line or box that already contains it. So the goal-directed step has nothing to offer, and we turn to the premise. ∀E at c gives $P(c) \wedge Q(c)$, and ∧E gives $P(c)$. The same works at d, at $f(c)$, at any term we name. But each line is an instance. No finite list of instances is the statement $\forall x\,P(x)$, which speaks about every object at once, including all the objects that none of our terms happen to name.

So with ∀E alone, a universal statement can be **used** but not **proved**. Leaving aside a detour through a contradiction, a derivation can write $\forall x\,\psi$ only by taking it out of a line or box that already contains it, as line 2 of (b) took $\forall y\,R(c, y)$ out of the premise. Section 1.3 was in the same position with implications, and Section 1.4 filled the gap with a box. Section 5.4 does the same for ∀: it proves P of one name about which nothing is known, and that box finishes this derivation in four lines.

Counter-readings. Each error above was shown wrong by a reading, a domain with a meaning for every symbol and a value for each free variable, in which the premises are true and the line is false. Call such a reading a **counter-reading**: it makes the whole counterexample set of the claim (Section 1.1), its premises and the negation of its line, true at once. A counter-reading shows that a step is unsafe. It is also how this book shows that a sequent has no derivation at all: a search that stalls proves nothing, since a cleverer derivation might exist. Sections 7.1 and 7.2 make readings exact and state that a counter-reading rules out every derivation. Until then, when this chapter says that something does not follow, it gives a counter-reading, and it never treats a stalled search as a proof.

5.3.8 The rule card

Example. Instances of ∀E

φ	t	*the instance* $\varphi[t/x]$
$P(x) \rightarrow Q(f(x))$	$f(f(c))$	$P(f(f(c))) \rightarrow Q(f(f(f(c))))$

$\exists\delta\,(\delta>0\wedge\forall y\,(\lvert y-x\rvert<\delta\to\lvert g(y)-g(x)\rvert<1))$	$a+1$	$\exists\delta\,(\delta>0\wedge\forall y\,(\lvert y-(a+1)\rvert<\delta\to\lvert g(y)-g(a+1)\rvert<1))$
$x^2\geq 0$	$\varepsilon/2$	$(\varepsilon/2)^2\geq 0$

In each row every free x gets the same term, and no letter of the term is caught (in the second row the quantifiers are on δ and y, and $a+1$ contains neither).

Not an instance. From $\forall x\,\exists y\,(x<y)$ at the term $y+1$, the line $\exists y\,(y+1<y)$ is not written by $\forall$E: the y of $y+1$ is caught by $\exists y$ (clause 4). Rename first; the instance is $\exists z\,(y+1<z)$, which says what was meant.

Quick check. Line 1 is $\forall x\,(x<f(x)\vee f(x)=x\cdot x)$. Is $2<f(2)\vee f(x)=x\cdot x$ an instance of it? Which clause decides?

Solution. No, clause 3: every free x gets the same term, so the instance at 2 is $2<f(2)\vee f(2)=2\cdot 2$.

Chapters 1 and 2's thirteen rules are on the card at the end of Section 3.1; here are this chapter's so far.

Rule	Shape	Side condition	You write	Lean
$\forall$E	$\forall x\,\varphi \Longrightarrow \varphi[t/x]$	t free for x in φ	"In particular, for $x=t$, …"; "applying the hypothesis to t, …"	`h t` (or `specialize h t`)

Written mathematics has many forms of the phrase: "in particular, for $x=3$, …"; "applying the hypothesis to x_0-1, …"; "by the theorem with $\varepsilon/2$ in place of ε, …". In the Lean column, `h` is the universal statement and `t` the term; applying `h` to `t` gives the instance, just as applying an implication to a proof of its hypothesis gave its conclusion in Section 1.3. In Lean a universal statement behaves like a function that takes a term and returns a proof of the instance.

Warning (Not allowed: $\forall$E). Each step below is wrong; the tag names the numbered clause of Definition 5.3.1 that it breaks.

- From $\forall x\,(x>0)\to Q(c)$, write $2>0\to Q(c)$: clause 1 (the main symbol is $\to$; the $\forall$ governs only $x>0$).
- From $\forall x\,R(x,f(x))$, write $R(c,f(d))$: clause 3 (one term in every free occurrence).
- From $\forall x\,\exists z\,(z\neq x)$, write $\exists z\,(z\neq z)$: clause 4 (the term z is captured).
- From $\forall x\in A\;\varphi$, write $\varphi[t/x]$ at once: clause 3 (the new line is the whole scope with t put in; the main symbol is $\forall x$, so clause 1 holds). The formula is $\forall x\,(x\in A\to\varphi)$, so $\forall$E gives $t\in A\to\varphi[t/x]$, and $\to$E needs $t\in A$.

5.3.9 The search method: one more step

The search method of Section 3.1 had a goal-directed step and a premise-directed step for each connective. Instantiation adds a premise-directed step.

- **A resource $\forall x\,\varphi$.** Do not instantiate blindly. Instantiate at the term that the goal, or another resource, asks for: match the shape of the goal (or of a hypothesis still needed) against φ, and read off the term. If the term is not free for x in φ, rename the bound variables of φ first.

In the Givens/Goal table, $\forall$E adds an instance $\varphi[t/x]$ to the Givens, with t read off the goal.

Instantiating is always **safe**, in the sense of Section 3.1: the new line is added and nothing is lost, since the universal line stays available. What can go wrong is only that the derivation fills with instances nobody needs, as a forward search filled with useless lines in Section 1.3.

5.3.10 Exercises

5.3.10.1 A. Check your understanding

Exercise. A1

(a) Write the schema of $\forall$E and its side condition from memory.

(b) For each pair, decide whether line 2 follows from line 1 by one application of $\forall$E. If not, name the clause of Definition 5.3.1 that fails.

- (i) Line 1: $\forall x\,\exists y\,(y > x)$; line 2: $\exists y\,(y > y)$.
- (i) Line 1: $\forall x\,(x \cdot x \geq 0)$; line 2: $3 \cdot x \geq 0$.
- (i) Line 1: $\forall x\,P(x) \to P(c)$; line 2: $P(a) \to P(c)$.
- (i) Line 1: $\forall x\,(P(x) \to P(f(x)))$; line 2: $P(f(c)) \to P(f(f(c)))$.

(d) True or false, with a reason: from $\forall x\,\forall y\,R(x,y)$, one application of $\forall$E gives $\forall x\,R(x,c)$.

(e) Why does $\forall$E not need a line saying that the term t names an element of the domain?

Solution.

(a) From line i: $\forall x\,\varphi$, line k may be $\varphi[t/x]$, justified "$\forall$E i", for any term t that is free for x in φ, with line i available at line k.

(b) (i) Not legal: the term y is not free for x in $\exists y\,(y > x)$, since the free x lies in the scope of $\exists y$ (clause 4, capture). Over $\mathbb{Z}$, line 1 is true and line 2 is false.

(ii) Not legal: only the first x was replaced, so line 2 is not $\varphi[t/x]$ for any t (clause 3). The instance at 3 is $3 \cdot 3 \geq 0$. Line 2 still has x free, and over $\mathbb{R}$ it is false when x names -1.

(iii) Not legal: the main symbol of line 1 is $\to$, not $\forall$ (clause 1); the $\forall x$ governs only $P(x)$. Over $\mathbb{N}$ with $P(x)$ read as $x \geq 1$, c as 0, and a as 1: line 1 says "if every natural number is at least 1, then $0 \geq 1$", true because its hypothesis is false; line 2 says "if $1 \geq 1$, then $0 \geq 1$", which is false.

(iv) Legal: it is $\varphi[f(c)/x]$ with $\varphi = P(x) \to P(f(x))$, both occurrences of x replaced by $f(c)$, and $f(c)$ has no variables, so it is free for x (clause 2: any term, compound ones included).

(v) False. The main symbol is $\forall x$, so one application removes $\forall x$ and gives $\forall y\,R(t,y)$ for

some term t. The line $\forall x\, R(x, c)$ would need the inner quantifier removed while the outer one stays, and the rule cannot reach inside the outer quantifier. (The sequent $\forall x\, \forall y\, R(x, y) \vdash \forall x\, R(x, c)$ is derivable, but it needs the rule of Section 5.4.)

(vi) Every term names an element of the domain: once the symbols are interpreted, a constant names one, and a compound term names the value of a function at elements. And the domain is never empty (Section 5.1). So "every element satisfies φ" applies to whatever t names, with nothing further to check. (Bounded quantifiers are different: to use $\forall x \in A\ \varphi$ at t one also needs $t \in A$, and Section 6.1 builds that into a rule.)

Exercise. A2: Drill

Each part takes a line or two. The parts run from single steps, through a missing premise, a justification column and a judgement on one step, to whole derivations.

(a) Line 1 is $\forall x\, (x^2 \geq 0)$. Write its instance at $a + 1$.

(b) Line 1 is $\forall x\, (P(x) \to Q(f(x)))$. What is its instance at $f(c)$?

(c) Line 1 is $\forall x\, \forall y\, (x + y = y + x)$. What does one $\forall$E at 2 give?

(d) Line 4 is $P(g(c)) \to Q(c)$, justified "$\forall$E 1". Give two different formulas that line 1 can be.

(e) Justify each line after the premises: 1 $\forall x\, (P(x) \to R(x, f(x)))$, premise; 2 $P(c)$, premise; 3 $P(c) \to R(c, f(c))$; 4 $R(c, f(c))$; 5 $R(c, f(c)) \wedge P(c)$.

(f) Line 1 is $\forall x\, P(x) \wedge Q(c)$. May $\forall$E cite it? If not, which clause does it break?

(g) Line 1 is $\forall x\, R(x, x)$. Is $R(c, d)$ an instance? If not, which clause does it break?

(h) Line 1 is $\forall x\, \exists y\, (y > x)$. Is $\forall$E at the term y legal? If not, which clause does it break?

(i) Same line: is $\forall$E at $x + 1$ legal, and what does it give?

(j) Derive $\forall x\, (P(x) \wedge Q(x)) \vdash Q(a)$.

(k) Derive $\forall x\, (P(x) \to Q(x)),\ \forall x\, P(x) \vdash Q(c)$.

(l) Derive $\forall x\, \forall y\, R(x, y) \vdash R(a, b) \wedge R(b, a)$.

Solution.

(a) $(a + 1)^2 \geq 0$.

(b) $P(f(c)) \to Q(f(f(c)))$.

(c) $\forall y\, (2 + y = y + 2)$: the outer quantifier only.

(d) Any two of $\forall x\, (P(g(x)) \to Q(x))$ (at c), $\forall x\, (P(x) \to Q(c))$ (at $g(c)$), $\forall x\, (P(g(x)) \to Q(c))$ (at c), $\forall x\, (P(g(c)) \to Q(x))$ (at c); also the degenerate $\forall x\, (P(g(c)) \to Q(c))$, with no free x.

(e) 3 $\forall$E 1 (at c); 4 $\to$E 3, 2; 5 $\wedge$I 4, 2.

(f) No, clause 1: its main connective is $\wedge$; $\wedge$E first gives $\forall x\, P(x)$.

(g) No, clause 3: every free x gets the same term, so the instances are $R(t, t)$.

(h) No, clause 4: y would be captured by $\exists y$. Rename first, and get $\exists z\, (z > y)$.

(i) Yes: $\exists y\, (y > x + 1)$. The only quantifier inside is on y, which is not in $x + 1$, so clause 4 holds.

(j) 2 $P(a) \wedge Q(a)$, $\forall$E 1; 3 $Q(a)$, $\wedge$E 2.

(k) 3 $P(c) \to Q(c)$, $\forall$E 1; 4 $P(c)$, $\forall$E 2; 5 $Q(c)$, $\to$E 3, 4.

(l) 2 $\forall y\, R(a, y)$, $\forall$E 1; 3 $R(a, b)$, $\forall$E 2; 4 $\forall y\, R(b, y)$, $\forall$E 1; 5 $R(b, a)$, $\forall$E 4; 6 $R(a, b) \wedge R(b, a)$, $\wedge$I 3, 5.

5.3.10.2 B. Practice

Exercise. B1: Two links

Give a derivation of

$$\forall x\,(P(x) \to Q(x)),\ \forall x\,(Q(x) \to R(x)),\ P(a) \vdash R(a),$$

and say, for each $\forall$E line, which line of the search asked for its term.

Solution. The goal $R(a)$ is the conclusion of the instance of the second premise at a, so it suffices to reach $Q(a)$; that is the conclusion of the instance of the first premise at a, so it suffices to reach $P(a)$, a premise. Both terms come from the goal: the second $\forall$E (line 6) from $R(a)$, the first (line 4) from the subgoal $Q(a)$.

1	$\forall x\,(P(x) \to Q(x))$	
2	$\forall x\,(Q(x) \to R(x))$	
3	$P(a)$	
4	$P(a) \to Q(a)$	$\forall$E 1
5	$Q(a)$	$\to$E 4, 3
6	$Q(a) \to R(a)$	$\forall$E 2
7	$R(a)$	$\to$E 6, 5

Exercise. B2: An increasing function

Read f as a function from $\mathbb{R}$ to $\mathbb{R}$ and a, b, c as real numbers. Give a derivation of

$$\forall x\,\forall y\,(x \le y \to f(x) \le f(y)),\ a \le b,\ b \le c \ \vdash\ f(a) \le f(b) \wedge f(b) \le f(c),$$

and then write the argument as one sentence of prose.

Solution. The goal is a conjunction, so its last line is $\wedge$I, with subgoals $f(a) \le f(b)$ and $f(b) \le f(c)$. Each is the conclusion of an instance of the first premise: matching $f(x) \le f(y)$ with $f(a) \le f(b)$ gives $x := a$, $y := b$, and with $f(b) \le f(c)$ gives $x := b$, $y := c$. The hypotheses of the two instances are the other two premises.

1	$\forall x\, \forall y\, (x \leq y \to f(x) \leq f(y))$	
2	$a \leq b$	
3	$b \leq c$	
4	$\forall y\, (a \leq y \to f(a) \leq f(y))$	$\forall$E 1
5	$a \leq b \to f(a) \leq f(b)$	$\forall$E 4
6	$f(a) \leq f(b)$	$\to$E 5, 2
7	$\forall y\, (b \leq y \to f(b) \leq f(y))$	$\forall$E 1
8	$b \leq c \to f(b) \leq f(c)$	$\forall$E 7
9	$f(b) \leq f(c)$	$\to$E 8, 3
10	$f(a) \leq f(b) \wedge f(b) \leq f(c)$	$\wedge$I 6, 9

The premise is instantiated twice, along two different paths: once at a and b, once at b and c. In prose: *Since f is increasing and $a \leq b$, we have $f(a) \leq f(b)$, and likewise $b \leq c$ gives $f(b) \leq f(c)$.* The word “increasing” carries the first premise, and the two applications of it are the two pairs of $\forall$E lines.

Exercise. B3: Squares are not negative

Let a and b be real numbers. Give a derivation of

$$\forall x\, (x^2 \geq 0) \vdash (a+b)^2 \geq 0 \wedge (a-b)^2 \geq 0.$$

Then extend it by one line justified “arith” to a derivation of $a^2 + b^2 \geq 2ab$ from the same premise, and say which earlier line the arith line uses.

Solution. The goal is a conjunction of two instances of the premise, at the compound terms $a+b$ and $a-b$. Both terms are free for x in $x^2 \geq 0$, which has no quantifier.

1	$\forall x\, (x^2 \geq 0)$	
2	$(a+b)^2 \geq 0$	$\forall$E 1
3	$(a-b)^2 \geq 0$	$\forall$E 1
4	$(a+b)^2 \geq 0 \wedge (a-b)^2 \geq 0$	$\wedge$I 2, 3
5	$a^2 + b^2 \geq 2 \cdot a \cdot b$	arith 3

Lines 1–4 are the first derivation. For the second goal only line 3 is needed: $(a-b)^2 = a^2 - 2ab + b^2$, so $(a-b)^2 \geq 0$ says $a^2 - 2ab + b^2 \geq 0$, and adding $2ab$ to both sides gives $a^2 + b^2 \geq 2ab$. That algebra is the arith line 5. (Lines 2 and 4 are then unused, which is allowed, and deleting them leaves a three-line derivation of $a^2 + b^2 \geq 2ab$. Line 2 would give $a^2 + b^2 \geq -2ab$ in the same way.)

Exercise. B4: Read the terms off the goal

A search for $\forall x\,(P(x) \to Q(x)),\ \forall x\,(Q(x) \to R(x)),\ \forall x\,P(x) \vdash R(c) \wedge Q(c)$ has reached the stage below. *Name the move, then complete it.*

1	$\forall x\,(P(x) \to Q(x))$	
2	$\forall x\,(Q(x) \to R(x))$	
3	$\forall x\,P(x)$	
	$\vdots$	
9	$R(c) \wedge Q(c)$	?

Solution. The move is *instantiate at the term you need*: the goal names c, so every $\forall$E is at c. $\wedge$I splits the goal into $R(c)$ and $Q(c)$, and each is the end of a chain of $\to$E.

1	$\forall x\,(P(x) \to Q(x))$	
2	$\forall x\,(Q(x) \to R(x))$	
3	$\forall x\,P(x)$	
4	$P(c)$	$\forall$E 3
5	$P(c) \to Q(c)$	$\forall$E 1
6	$Q(c)$	$\to$E 5, 4
7	$Q(c) \to R(c)$	$\forall$E 2
8	$R(c)$	$\to$E 7, 6
9	$R(c) \wedge Q(c)$	$\wedge$I 8, 6

Exercise. B5: Derive or break

For each claim, either derive it with the rules of Chapters 1–2 and $\forall$E, or break it with a counter-reading: a domain, a meaning for each symbol, and a value for each constant, that make its counterexample set true. Then pair each of the breaks (e) and (g) with the derivable claim in the list that it imitates, and say what separates them.

(a) $\forall x\,(P(x) \vee Q(x)),\ \neg P(c) \vdash Q(c)$.
(b) $\forall x\,(P(x) \to Q(x)),\ Q(c) \vdash P(c)$.
(c) $\forall x\,R(x, f(x)) \vdash R(c, f(c)) \wedge R(f(c), f(f(c)))$.
(d) $\forall x\,R(x, f(x)) \vdash R(f(c), c)$.
(e) $\forall x\,(P(x) \to \neg P(f(x))),\ P(c) \vdash \neg P(f(f(c)))$.
(f) $\forall x\,(P(x) \to \neg P(f(x))),\ P(f(c)) \vdash \neg P(c) \wedge \neg P(f(f(c)))$.
(g) $\forall x\,\exists y\,R(x, y) \vdash \exists y\,R(y, y)$.
(h) $\forall x\,\exists y\,R(x, y) \vdash \exists y\,R(c, y)$.

Solution.

(a) Derive. Instantiate at c, then take the disjunction apart by cases.

1	$\forall x\,(P(x) \lor Q(x))$	
2	$\neg P(c)$	
3	$P(c) \lor Q(c)$	$\forall$E 1
4	$\quad P(c)$	asm, for $\lor$E
5	$\quad \bot$	$\neg$E 4, 2
6	$\quad Q(c)$	$\bot$E 5
7	$\quad Q(c)$	asm, for $\lor$E
8	$Q(c)$	$\lor$E 3, 4–6, 7–7

(b) Break. The domain $\{0\}$, c naming 0, P true of nothing, Q true of 0. The first premise holds (its hypothesis is false at 0), $Q(c)$ holds, and $P(c)$ fails. (The form affirms the consequent of an instance.)

(c) Derive. Two instances of the one premise, at c and at $f(c)$.

1	$\forall x\,R(x, f(x))$	
2	$R(c, f(c))$	$\forall$E 1
3	$R(f(c), f(f(c)))$	$\forall$E 1
4	$R(c, f(c)) \land R(f(c), f(f(c)))$	$\land$I 2, 3

(d) Break. The domain $\{0, 1\}$, c naming 0, $f(0) = 1$ and $f(1) = 1$, R true of the pairs $(0, 1)$ and $(1, 1)$ only. The premise holds: $R(0, f(0))$ is $R(0, 1)$, and $R(1, f(1))$ is $R(1, 1)$. The conclusion $R(f(c), c)$ is $R(1, 0)$, which is false.

(e) Break. The domain $\{0, 1\}$, c naming 0, f exchanging 0 and 1, P true of 0 only. The first premise holds: at 0, $P(0)$ holds and $P(1)$ does not; at 1, the hypothesis $P(1)$ is false. $P(c)$ holds, and $\neg P(f(f(c)))$ is $\neg P(0)$, which is false.

(f) Derive. The instance at $f(c)$ gives the right conjunct at once; the instance at c refutes $P(c)$ inside a $\neg$I box.

1	$\forall x\,(P(x) \to \neg P(f(x)))$	
2	$P(f(c))$	
3	$P(f(c)) \to \neg P(f(f(c)))$	$\forall$E 1
4	$\neg P(f(f(c)))$	$\to$E 3, 2
5	$P(c) \to \neg P(f(c))$	$\forall$E 1
6	$\quad P(c)$	asm, for $\neg$I
7	$\quad \neg P(f(c))$	$\to$E 5, 6
8	$\quad \bot$	$\neg$E 2, 7
9	$\neg P(c)$	$\neg$I 6–8
10	$\neg P(c) \land \neg P(f(f(c)))$	$\land$I 9, 4

(g) Break. The domain $\{0,1\}$, R true of $(0,1)$ and $(1,0)$ only ("differs from"). Every element has a partner, and no element is its own.

(h) Derive, in one step: the only quantifier inside is on y, and c contains no y.

$$\begin{array}{lll} 1 & \forall x\,\exists y\,R(x,y) & \\ \hline 2 & \exists y\,R(c,y) & \forall\text{E } 1 \end{array}$$

(e) and (f) are look-alikes: the premise about f is used at the wrong term in (e), at the right one in (f). (g) is the capture look-alike of (h): instantiating at the bound letter y would write $\exists y\,R(y,y)$. (b) is the converse look-alike of the square-root step of Example (Instantiation at work) (a). Each break is a counter-reading of the kind in the figure of "Common errors".

Marking scheme. (10 points) - 1 for each of the eight verdicts with its derivation or counter-reading. A counter-reading must give the domain, every symbol's meaning and the value of c, and check each premise and the conclusion. - 2: the closing pairing, (e) with (f) and (g) with (h), with the reason. Fatal errors (0 for the claim): a "break" by a reading in which a premise is false; a $\forall$E at a term that is not free for its variable.

5.3.10.3 C. Going deeper

Exercise. C1: A partner makes a loop

A relation R is **symmetric** if $R(x,y)$ gives $R(y,x)$, and **transitive** if $R(x,y)$ and $R(y,z)$ give $R(x,z)$. Show that in a symmetric and transitive relation, every element that is related to something is related to itself, by giving a derivation of

$$\forall x\,\forall y\,(R(x,y) \to R(y,x)),$$
$$\forall x\,\forall y\,\forall z\,((R(x,y) \land R(y,z)) \to R(x,z)),\ R(a,b) \ \vdash\ R(a,a).$$

Explain how each term of each $\forall$E was chosen.

Solution (Hint 1). The goal $R(a,a)$ can only come from the transitivity premise. Match its conclusion $R(x,z)$ with the goal.

Solution (Hint 2). The match gives $x \coloneqq a$ and $z \coloneqq a$, and leaves y open. Which y makes the hypothesis $R(a,y) \land R(y,a)$ reachable from the premises?

Solution. The search. The goal $R(a,a)$ is atomic, so it must be the conclusion of an instance of a premise. Symmetry ends in $R(y,x)$, and its instance at a,a is $R(a,a) \to R(a,a)$, whose hypothesis is the goal itself: a dead end. Transitivity ends in $R(x,z)$. Matching gives $x \coloneqq a$ and $z \coloneqq a$. The variable y does not occur in the conclusion, so the goal says nothing about it; the hypothesis $R(a,y) \land R(y,a)$ must decide. The premise $R(a,b)$ offers $R(a,y)$ with $y \coloneqq b$. Then the second conjunct is $R(b,a)$, which symmetry gives from $R(a,b)$: the instance of the symmetry premise at $x \coloneqq a$, $y \coloneqq b$ (read off its needed conclusion $R(b,a)$, as in Example (Instantiation at work) (c)).

1	$\forall x\,\forall y\,(R(x,y) \to R(y,x))$	
2	$\forall x\,\forall y\,\forall z\,((R(x,y) \land R(y,z)) \to R(x,z))$	
3	$R(a,b)$	
4	$\forall y\,(R(a,y) \to R(y,a))$	$\forall$E 1
5	$R(a,b) \to R(b,a)$	$\forall$E 4
6	$R(b,a)$	$\to$E 5, 3
7	$\forall y\,\forall z\,((R(a,y) \land R(y,z)) \to R(a,z))$	$\forall$E 2
8	$\forall z\,((R(a,b) \land R(b,z)) \to R(a,z))$	$\forall$E 7
9	$(R(a,b) \land R(b,a)) \to R(a,a)$	$\forall$E 8
10	$R(a,b) \land R(b,a)$	$\land$I 3, 6
11	$R(a,a)$	$\to$E 9, 10

The terms. Lines 7 and 9 put a in for x and for z: both read off the goal. Line 8 puts b in for y: read off the premise $R(a,b)$, the only available fact of the form $R(a,\cdot)$. Lines 4 and 5 put in a and b: read off the needed line $R(b,a)$. All terms are constants, so every one is free for its variable.

The hypothesis “related to something” is needed. In a symmetric and transitive relation an element related to nothing need not be related to itself: on $\{0,1\}$, the relation that holds only between 0 and 0 is symmetric and transitive, and $R(1,1)$ fails. Without a premise like $R(a,b)$ the search has no term for y that makes the hypothesis of line 9 reachable.

Marking scheme. (10 points) - 2: identifies transitivity as the source of the goal and instantiates it with a for x and z. - 2: chooses b for y, with the reason (the premise $R(a,b)$). - 2: obtains $R(b,a)$ from symmetry at a, b, with $\to$E. - 2: builds the conjunction in the right order and fires the transitivity instance. - 2: every $\forall$E removes the outermost quantifier of a cited line and replaces every occurrence of its variable. Fatal errors (0 for the step): removing an inner quantifier before an outer one; instantiating transitivity at a term for y for which neither conjunct of the hypothesis is available.

Exercise. C2: Instantiating at the variable itself

(a) Show that $\forall x\,P(x) \vdash P(x)$, by a derivation whose one $\forall$E line you check against every clause of Definition 5.3.1.

(b) The reverse step, from a line $P(x)$ to $\forall x\,P(x)$, is an instance of no rule so far. Show that it must not be allowed without restriction: give a reading over $\mathbb{Z}$, with x naming one particular integer, in which $P(x)$ is true and $\forall x\,P(x)$ is false.

Solution.

(a)

1	$\forall x\,P(x)$	
2	$P(x)$	$\forall$E 1

Clause 1: the main symbol of line 1 is $\forall x$, with scope $P(x)$. Clause 2: the term is the

variable x itself, and any term is allowed. Clause 3: $P(x)[x/x] = P(x)$, so every free occurrence was replaced, by x. Clause 4: x is free for x in every formula (Section 5.2: a free occurrence of x never lies in the scope of a quantifier on x). Clause 5: line 1 is a premise and lies in no box, so it is available at line 2. So line 2 is legal. Its x is now free: line 2 says "P holds of whatever x names", which is true whatever x names, since line 1 says that P holds of everything.

(b) Read $P(x)$ as "x is even" and let x name 2. Then $P(x)$ is true, and $\forall x\, P(x)$, "every integer is even", is false, since 1 is odd. A rule allowing the step from $P(x)$ to $\forall x\, P(x)$ in general would lead from a true line to a false one.

Marking scheme. (10 points) - 6: (a) the derivation, with each of the five numbered clauses checked (main symbol $\forall x$; any term; every free occurrence by the same term; free for x; availability), including that x is free for x. - 4: (b) a reading with $P(x)$ true and $\forall x\, P(x)$ false, with the value of x stated. Fatal error (0 for the step): in (b), a reading in which $P(x)$ is false or x has no value.

5.4 Proving "For Every": Let x Be Arbitrary

Section 5.3 gave the rule for using a universal statement, and ended at a stall: from $\forall x\,(P(x) \wedge Q(x))$ it could derive $P(c)$ for every term c we cared to name, but never $\forall x\, P(x)$ itself. No list of instances is a statement about everything. This section gives the rule that proves a universal statement, $\forall$I. It needs one new device, a box that carries a **name** instead of an assumption, and one new side condition, which turns the mathematician's "let x be arbitrary" into something a reader can check line by line.

5.4.1 One number we know nothing about

Here is a fact of school algebra.

> For every real number x, $x^2 - 2x + 1 \geq 0$.

We cannot check it number by number; there are too many numbers. The proof everyone writes takes **one** real number and argues about it: *Let $x \in \mathbb{R}$. Then $x^2 - 2x + 1 = (x-1)^2$, and a square is never negative. So $x^2 - 2x + 1 \geq 0$.* Then, often without saying so: *since x was arbitrary, this holds for every real x.*

Chapter 1 adopted a convention for sentences like "let x be a real number": the letter names **one fixed** number, not further specified (Section 1.1). It promised that Chapter 5 would explain why an argument of that kind proves a statement about every x at once. Section 2.2 met the same step at the end of the proof of Theorem 2.2.1, passing from an arbitrary rational r to every rational; Section 4.2 met it in induction, passing from "let $n \in \mathbb{N}$" to "for every n". Each time the step was taken on trust. Here is why it is correct, and when it is not.

The argument about x used **nothing** about x except that it is a real number. It never said "suppose $x > 0$" or "take $x = 3$". So it could be repeated word for word with 0, with -7, with π, with any real number in place of x, and it would prove the claim for that number. One argument about a number we know nothing about is a recipe for an argument about each number, and that is what "for every" asks for.

Everything depends on "nothing". If the argument had used a fact about x that not every

number has, the recipe would break for some numbers. "Let x be a real number with $x > 0$. Then $x > 0$. Since x was arbitrary, every real number is positive" is the same shape, and its conclusion is false. The step "since x was arbitrary" is legitimate only when x **was** arbitrary: when no premise and no assumption still in force says anything about it. The rule of this section makes that condition exact.

5.4.2 The flagged box

To prove "for every x", take a fresh name that nothing is known about, prove the statement for that name, and generalize.

In a derivation the argument about the arbitrary x is drawn as a box, as Chapter 1 drew the argument under an assumption. The box opens with no assumption at all. Instead it carries the arbitrary name, written at its top left, and is called a **flagged box**; the name is its **flag**.

Definition 5.4.1. Rules for the Universal Quantifier

$$\begin{array}{ll|l|l} i & x_0 & \vdots & \\ j & & \varphi[x_0/x] & \\ k & \forall x\,\varphi & & \forall\text{I } i\text{–}j \end{array} \qquad \begin{array}{ll|l|l} i & \forall x\,\varphi & & \\ k & \varphi[t/x] & & \forall\text{E } i \end{array}$$

The rule on the left is ∀I (**universal introduction**, or **generalization**). It cites a flagged box, with flag x_0, from its first line i to its last line j, and writes $\forall x\,\varphi$ on line k. The flag x_0 is a **variable**. The box's last line j is $\varphi[x_0/x]$, and it is a line of the flagged box **itself**, not of a box inside it. The side condition is on the name: at line k,

- x_0 appears in **no premise**,
- x_0 appears in **no assumption of a box still open at line** k, and
- x_0 does **not** appear in $\forall x\,\varphi$.

The rule on the right is ∀E, as in Definition 5.3.1: t is any term free for x in φ.

In words: to prove that φ holds for every x, open a box flagged with a new name x_0, prove $\varphi[x_0/x]$, "φ holds of x_0", as the box's last line, close the box, and write $\forall x\,\varphi$. The side condition says that x_0 really was arbitrary.

5.4.2.1 A first flagged box: how the rule looks in a mathematical proof

Here is the opening proof as a derivation, with the general fact about squares as a premise and the algebra as one line justified "arith" (Section 5.3).

$$\begin{array}{ll|l} 1 & & \forall y\,(y^2 \geq 0) \\ \hline & x_0 & \vdots \\ 3 & & x_0^2 - 2\cdot x_0 + 1 \geq 0 \\ 4 & & \forall x\,(x^2 - 2\cdot x + 1 \geq 0) \end{array}$$

The goal is universal, so the box opens first, with the goal's instance at x_0 as its last line; then the premise is instantiated at the term that last line needs, $x_0 - 1$.

1	$\forall y\,(y^2 \geq 0)$	
2	x_0 | $(x_0 - 1)^2 \geq 0$	$\forall$E 1
	$\vdots$	
3	$x_0^2 - 2 \cdot x_0 + 1 \geq 0$	
4	$\forall x\,(x^2 - 2 \cdot x + 1 \geq 0)$	

The finished derivation:

1	$\forall y\,(y^2 \geq 0)$	
2	x_0 | $(x_0 - 1)^2 \geq 0$	$\forall$E 1
3	$x_0^2 - 2 \cdot x_0 + 1 \geq 0$	arith 2
4	$\forall x\,(x^2 - 2 \cdot x + 1 \geq 0)$	$\forall$I 2–3

Line	Sentence
box x_0	"Let $x \in \mathbb{R}$."
2	"Then $(x-1)^2 \geq 0$, since a square is never negative." ($\forall$E at $x_0 - 1$)
3	"That is, $x^2 - 2x + 1 \geq 0$." (arith)
4	"Since x was arbitrary, $x^2 - 2x + 1 \geq 0$ for every real x." ($\forall$I)

Lines 2 and 3 are the derivation of Section 5.3 with the flag x_0 in place of the constant a. What changed is only the name: a was a constant that the language fixed, so nothing could be generalized over it; x_0 is a variable that no premise mentions, so line 4 may generalize. The written proof uses the same letter x for the flag and for the bound variable, and says "let $x \in \mathbb{R}$" where the derivation opens the box. Readers of mathematics are expected to supply the box.

Since the domain is $\mathbb{R}$, "let $x \in \mathbb{R}$" asks nothing of x. A proof that says "let $\varepsilon > 0$" does ask something: inside the box flagged ε_0, a $\to$I box opens with the assumption $\varepsilon_0 > 0$.

5.4.2.2 Reading the rule

Here are the two boxes side by side: the box of $\to$I from Definition 1.4.1, and the flagged box.

i	| φ	asm, for $\to$I
	| $\vdots$	
j	| ψ	
k	$\varphi \to \psi$	$\to$I i–j

i	x_0 | $\vdots$	
j	| $\varphi[x_0/x]$	
k	$\forall x\,\varphi$	$\forall$I i–j

Both draw a boundary and both are cited as a range i–j. The $\to$I box opens with what we suppose; the flagged box opens with the name we have fixed. In the schema the flagged box begins with a gap $\vdots$ on line i, only so that the range has a first line; in a real derivation line i is whatever the box's first line is.

This rule has conditions that matter; read each one. Read the rule part by part: the flag, the three clauses of the side condition, and the last line.

- **The flag.** Opening the box is the written "let x_0 be arbitrary". The flag is always a variable, never a constant: a constant has a fixed meaning, while the whole point of x_0 is that it may stand for anything. The book writes flags with a subscript 0, x_0, y_0, ε_0, so that they stand out.
- **No premise, no open assumption** (the first and second clauses). This is "we know nothing about x_0". What is known at line k is the premises and the assumptions of the boxes still open there. If none of them mentions x_0, then nothing that was used about x_0 inside the box is special to x_0.
- **Not in the conclusion** (the third clause). The name must not leak out. Inside the box it stood for an arbitrary object, and $\forall x\,\varphi$ is meant to speak about all objects through the bound x. A conclusion that still mentions x_0 ties what it says to the one object x_0 names, while claiming it of everything.
- **The last line.** $\varphi[x_0/x]$ is what was proved about x_0, and it must be a line of the flagged box itself, just as the last line of a $\to$I box must be (Section 1.4). A line inside a further box would still depend on that box's assumption.

"Appear" is the word of Section 5.2: written anywhere in the formula, including just after a quantifier symbol. So a box flagged x can never be closed by $\forall$I with $\forall x\,\varphi$, since x appears in $\forall x\,\varphi$. That is one more reason the book never uses the quantified letter itself as the flag.

In practice: choose as flag a letter written nowhere in the derivation so far, and as the quantified letter x one that does not appear in the box's last line (if it does, rename that bound letter first, Section 5.2). When the box closes, put x in place of **every** occurrence of the flag in its last line. Then the last line is $\varphi[x_0/x]$ for the formula φ so obtained, and all three clauses hold at once: no premise or open assumption can contain a letter written nowhere, and the conclusion no longer contains the flag. The clauses are what a checker tests, and they allow a little more than this habit does (the fold below).

↻ **Second reading** · *Fresh is not the same as new*

The condition asks about three places only: the premises, the open assumptions, and the conclusion. Lines **outside** the box may mention x_0, before it or after it. An earlier line $\forall y\,R(x_0, y)$, obtained by $\forall$E from a premise, is harmless; Exercise C1 builds a derivation with such a line. And once a flagged box has closed, its name may flag a later box. A name that satisfies the side condition at the line that closes its box is called **fresh** there. It need not be new to the whole derivation.

Quick check. A derivation has used the letters x_0, y_0 and u_0 so far, and u_0 appears in an open assumption. Which flag can a new $\forall$I box take without checking anything?

Solution. Any letter written nowhere yet, for example z_0. (x_0 or y_0 may also be allowed, but only after checking the three clauses; u_0 is not, while its assumption is open.)

5.4.2.3 Why each clause

Each clause stops a derivation of a false conclusion from true premises. Here is one failure for each.

A name in a premise. If ∀I could generalize over a name in a premise, then $P(x_0) \vdash \forall x\, P(x)$:

1	$P(x_0)$	
2	x_0 | $P(x_0)$	R 1
3	$\forall x\, P(x)$	∀I 2–2

Over $\mathbb{Z}$, read $P(x)$ as "x is even" and let x_0 name 2. The premise is true and line 3, "every integer is even", is false. Line 3 breaks the first clause: x_0 appears in the premise on line 1. The premise **does** say something about x_0, so x_0 is not arbitrary.

A name in an open assumption. Suppose $x_0 > 0$, and inside that supposition generalize over x_0 (left):

1	|	$x_0 > 0$	asm, for →I
2	|	x_0 | $x_0 > 0$	R 1
3	|	$\forall x\,(x > 0)$	∀I 2–2
4	$x_0 > 0 \to \forall x\,(x > 0)$		→I 1–3

1	x_0 |	|	$x_0 > 0$	asm, for →I
2	|	|	$x_0 + 1 > 0$	arith 1
3	|	$x_0 > 0 \to x_0 + 1 > 0$		→I 1–2
4	$\forall x\,(x > 0 \to x + 1 > 0)$			∀I 1–3

On the left, line 3 breaks the second clause, since x_0 appears in the assumption of line 1, whose box is still open at line 3. Read over $\mathbb{Z}$ with x_0 naming 1: line 4 says "if $1 > 0$, then every integer is positive", which is false at 0, and it rests on no premise at all. This is the "let x be a real number with $x > 0$ … since x was arbitrary" of the opening, made into lines. One more legal flagged box around everything turns line 4 into a false sentence; the fold shows it. On the right the boxes are nested the other way: the flagged box opens first, the supposition inside it, and ∀I generalizes an implication that carries its hypothesis with it.

↻ **Second reading** · *The same error, as a false sentence*

With one more flagged box, the error of the minimal derivation becomes a false **sentence**:

1	x_0 |	|	$x_0 > 0$	asm, for →I
2	|	|	x_0 | $x_0 > 0$	R 1
3	|	|	$\forall x\,(x > 0)$	∀I 2–2
4	|	$x_0 > 0 \to \forall x\,(x > 0)$		→I 1–3
5	$\forall y\,(y > 0 \to \forall x\,(x > 0))$			∀I 1–4

Line 3 breaks the second clause: x_0 appears in the assumption on line 1, whose box is still open at line 3. Every other line is legal, including line 5, where the box of line 1 has closed. The inner box on line 2 is flagged x_0 while the outer box flagged x_0 is still open; the definition allows this, since its side condition asks only about premises, open assumptions and the conclusion. And

line 5 says "every positive number makes every number positive", which over $\mathbb{Z}$ is false: at $y = 1$ the hypothesis $1 > 0$ holds and $\forall x\,(x > 0)$ fails at $x = 0$. This is the "let x be a real number with $x > 0 \ldots$ since x was arbitrary" of the opening, made into lines.

A name in the conclusion. From the premise $\forall x\,(x \le x)$, true over $\mathbb{N}$, open a box flagged y_0:

1	$\forall x\,(x \le x)$	
2	y_0 | $y_0 \le y_0$	$\forall$E 1
3	$\forall x\,(x \le y_0)$	$\forall$I 2–2

Line 2 is $(x \le y_0)[y_0/x]$, so line 3 has the shape of $\forall$I with $\varphi = (x \le y_0)$. But y_0 appears in the conclusion, which breaks the third clause. Line 3 says "every natural number is at most y_0", and that is false whatever y_0 names: $y_0 + 1$ is a counterexample. (Exercise C1 of Section 5.5 turns this into a sentence, "there is a largest natural number".)

The requirement on the last line is the one Chapter 1 already imposed on every box rule, for the same reason.

5.4.3 Derivations with flagged boxes

The search method of Section 3.1 gets a new goal-directed step: when the goal is $\forall x\,\varphi$, open a flagged box with a fresh name x_0, and make $\varphi[x_0/x]$ the box's last line, the new goal. The examples below use it, each with an Idea.

Example. Flagged boxes at work

Give a derivation of each sequent.

(a) $\forall x\,\forall y\,R(x,y) \vdash \forall x\,R(x,x)$.
(b) $\forall x\,(P(x) \to Q(x)),\ \forall x\,P(x) \vdash \forall x\,Q(x)$.
(c) $\forall x\,\forall y\,R(x,y) \vdash \forall y\,\forall x\,R(x,y)$.

Solution.

(a) *Idea.* The goal begins with $\forall x$, so the goal-directed step opens a box flagged x_0, and the new goal is $R(x,x)[x_0/x] = R(x_0,x_0)$. That goal names the term for $\forall$E (Section 5.3's step): instantiate the premise at x_0, twice.

1	$\forall x\,\forall y\,R(x,y)$	
	x_0 | $\vdots$	
3	| $R(x_0,x_0)$	
4	$\forall x\,R(x,x)$	

1	$\forall x\,\forall y\,R(x,y)$	
2	x_0 | $\forall y\,R(x_0,y)$	$\forall$E 1
	| $\vdots$	
3	| $R(x_0,x_0)$	
4	$\forall x\,R(x,x)$	

1	$\forall x\, \forall y\, R(x, y)$		
2	x_0 | $\forall y\, R(x_0, y)$		$\forall$E 1
3	| $R(x_0, x_0)$		$\forall$E 2
4	$\forall x\, R(x, x)$		$\forall$I 2–3

The side condition at line 4: x_0 does not appear in the premise, no box is open at line 4, and x_0 does not appear in $\forall x\, R(x, x)$. Lines 2 and 3 instantiate at the variable x_0, which is free for x and for y, since the formulas contain no quantifier on x_0. Compare Section 5.3, (b), which reached only $R(c, c)$: there the term was a constant, a name fixed by the language, and nothing can be generalized over a constant; here it is a flag, a variable that no premise mentions.

(b) *Idea.* The goal is universal: a box flagged x_0, whose last line must be $Q(x_0)$. Both premises are universal; instantiate each at the term the box's goal names, x_0, and $\to$E finishes.

Givens	Goal
$\forall x\, (P(x) \to Q(x))$, $\forall x\, P(x)$	$\forall x\, Q(x)$
the same, and the fresh name x_0 ($\forall$I opens a box)	$Q(x_0)$
+ $P(x_0) \to Q(x_0)$, $P(x_0)$ ($\forall$E at x_0, twice)	$Q(x_0)$
+ $Q(x_0)$ ($\to$E)	$Q(x_0)$ ✓

1	$\forall x\,(P(x) \to Q(x))$	
2	$\forall x\,P(x)$	
	x_0 ⋮	
5	$Q(x_0)$	
6	$\forall x\,Q(x)$	

1	$\forall x\,(P(x) \to Q(x))$	
2	$\forall x\,P(x)$	
3	x_0 $P(x_0) \to Q(x_0)$	$\forall$E 1
4	$P(x_0)$	$\forall$E 2
	⋮	
5	$Q(x_0)$	
6	$\forall x\,Q(x)$	

1	$\forall x\,(P(x) \to Q(x))$	
2	$\forall x\,P(x)$	
3	x_0 $P(x_0) \to Q(x_0)$	$\forall$E 1
4	$P(x_0)$	$\forall$E 2
5	$Q(x_0)$	$\to$E 3, 4
6	$\forall x\,Q(x)$	$\forall$I 3–5

At line 6, x_0 appears in neither premise nor in $\forall x\,Q(x)$, and no box is open.

(c) *Idea.* Read the goal's quantifiers from the outside: $\forall y$ first, so open a box flagged y_0; the new goal $\forall x\,R(x, y_0)$ is again universal, so open a box flagged x_0 inside it; the innermost goal is $R(x_0, y_0)$. Two boxes, opened in the goal's order. Now the premise's quantifiers are removed in **its** order: $\forall$E at x_0, then at y_0.

1	$\forall x\,\forall y\,R(x, y)$	
	y_0 x_0 ⋮	
3	$R(x_0, y_0)$	
4	$\forall x\,R(x, y_0)$	
5	$\forall y\,\forall x\,R(x, y)$	

1	$\forall x\,\forall y\,R(x, y)$	
2	y_0 x_0 $\forall y\,R(x_0, y)$	$\forall$E 1
	⋮	
3	$R(x_0, y_0)$	
4	$\forall x\,R(x, y_0)$	
5	$\forall y\,\forall x\,R(x, y)$	

1	$\forall x\,\forall y\,R(x, y)$	
2	y_0 x_0 $\forall y\,R(x_0, y)$	$\forall$E 1
3	$R(x_0, y_0)$	$\forall$E 2
4	$\forall x\,R(x, y_0)$	$\forall$I 2–3
5	$\forall y\,\forall x\,R(x, y)$	$\forall$I 2–4

The box flagged y_0 immediately contains the box flagged x_0, so it has no line of its own before the inner box; following the convention for a flagged box that immediately contains another box (the fine print below), line 5 cites it as 2–4, from the first line of the inner box. Its last line is line 4, a line of the outer box itself. At line 4, x_0 is in no premise, the only open box (flagged y_0) has no assumption, and x_0 does not appear in $\forall x\,R(x, y_0)$. At line 5 the same holds for y_0.

Nothing in (c) used that the formula under the quantifiers is $R(x, y)$. For any formula ψ and different variables x and y, with flags x_0 and y_0 that are different from each other and from x and y and that do not appear in ψ, the same five lines derive $\forall x\,\forall y\,\psi \vdash \forall y\,\forall x\,\psi$: line 3 becomes $(\psi[x_0/x])[y_0/y]$, which equals $(\psi[y_0/y])[x_0/x]$, since the two substitutions put two different fresh names in for two different variables and so do not interfere. Exchanging the letters x and y gives the other direction. So **two adjacent universal quantifiers may always be swapped**: $\forall x\,\forall y\,\psi \dashv\vdash \forall y\,\forall x\,\psi$. Section 5.1 read this off the game, where the challenger makes both moves; here it is derived.

Quick check. Suppose premise 2 of (b) were $P(x_0)$ in place of $\forall x\,P(x)$. The box still reaches $Q(x_0)$, by ∀E 1 and →E. May line 6 be written?

Solution. No: x_0 appears in a premise, so the first clause fails. A counter-reading shows the step is unsafe: over $\mathbb{Z}$, with P and Q both read as "even" and x_0 naming 2, both premises are true and $\forall x\,Q(x)$ is false.

Quick check. Could both boxes of (c) be flagged x_0, the outer one standing for y?

Solution. No. The inner box would end in $R(x_0, x_0)$, and the ∀I line that closes it must be $\forall x\,R(x, x_0)$, the instance at x_0 of the outer goal; that line contains the inner flag x_0, so the third clause fails. Here the nested flagged boxes need two names.

↻ **Second reading** · *The front renaming, derived*

Section 5.2 took renaming of a bound variable as a convention and promised to derive it at the front of a formula. For a formula φ and variables y, z with z not appearing in φ, $\forall y\,\varphi \dashv\vdash \forall z\,\varphi[z/y]$:

1	$\forall y\,\varphi$			1	$\forall z\,\varphi[z/y]$		
2	z_0 | $\varphi[z_0/y]$	∀E 1		2	z_0 | $\varphi[z_0/y]$	∀E 1	
3	$\forall z\,\varphi[z/y]$	∀I 2–2		3	$\forall y\,\varphi$	∀I 2–2	

Take z_0 to be a variable that does not appear in φ and is different from y and z. Since z does not appear in φ, the formula $\varphi[z/y]$ is φ with each free y changed to z, and changing those z's on to z_0 gives the same result as changing each free y to z_0 at once: $(\varphi[z/y])[z_0/z] = \varphi[z_0/y]$. So in the left derivation, line 2 is a legal ∀E (z_0 does not appear in φ, so it is free for y), and it is the instance $(\varphi[z/y])[z_0/z]$ that the ∀I on line 3 needs. In the right derivation, line 2 is $(\varphi[z/y])[z_0/z]$, a legal ∀E for the same reason, and it is $\varphi[z_0/y]$, the instance that $\forall y\,\varphi$ needs. At line 3 of each, z_0 appears neither in the premise nor in the conclusion, and no box is open.

Each direction has three lines.

Now the sequent at which Section 5.3 stalled, together with its converse. "For every x, $P(x)$ and $Q(x)$" says the same as "$P(x)$ for every x, and $Q(x)$ for every x":

Proposition 5.4.1. Universal Quantifier and Conjunction

$$\forall x\,(P(x) \wedge Q(x)) \dashv\vdash \forall x\,P(x) \wedge \forall x\,Q(x).$$

Idea. ($\vdash$) The goal is a conjunction, so (Section 1.2) two subgoals, $\forall x\,P(x)$ and $\forall x\,Q(x)$. Each is universal, so each gets its own flagged box, and inside each, $\forall$E at the flag and $\wedge$E give the box's last line. ($\dashv$) The goal is universal, so one flagged box, with the goal $P(x_0) \wedge Q(x_0)$; take the premise apart by $\wedge$E and instantiate each half at x_0.

The search, for ($\vdash$):

1	$\forall x\,(P(x) \wedge Q(x))$	
	x_0 ⋮	
3	$P(x_0)$	
4	$\forall x\,P(x)$	
	x_0 ⋮	
6	$Q(x_0)$	
7	$\forall x\,Q(x)$	
8	$\forall x\,P(x) \wedge \forall x\,Q(x)$	$\wedge$I 4, 7

1	$\forall x\,(P(x) \wedge Q(x))$	
2	x_0 $P(x_0) \wedge Q(x_0)$	$\forall$E 1
	⋮	
3	$P(x_0)$	
4	$\forall x\,P(x)$	
5	x_0 $P(x_0) \wedge Q(x_0)$	$\forall$E 1
	⋮	
6	$Q(x_0)$	
7	$\forall x\,Q(x)$	
8	$\forall x\,P(x) \wedge \forall x\,Q(x)$	$\wedge$I 4, 7

For ($\dashv$):

1	$\forall x\,P(x) \wedge \forall x\,Q(x)$	
2	$\forall x\,P(x)$	$\wedge$E 1
3	$\forall x\,Q(x)$	$\wedge$E 1
	x_0 ⋮	
6	$P(x_0) \wedge Q(x_0)$	
7	$\forall x\,(P(x) \wedge Q(x))$	

1	$\forall x\,P(x) \wedge \forall x\,Q(x)$	
2	$\forall x\,P(x)$	$\wedge$E 1
3	$\forall x\,Q(x)$	$\wedge$E 1
4	x_0 $P(x_0)$	$\forall$E 2
5	$Q(x_0)$	$\forall$E 3
	⋮	
6	$P(x_0) \wedge Q(x_0)$	
7	$\forall x\,(P(x) \wedge Q(x))$	

Proof. ($\vdash$)

1		$\forall x\,(P(x) \wedge Q(x))$	
2	x_0	$P(x_0) \wedge Q(x_0)$	$\forall$E 1
3		$P(x_0)$	$\wedge$E 2
4		$\forall x\,P(x)$	$\forall$I 2–3
5	x_0	$P(x_0) \wedge Q(x_0)$	$\forall$E 1
6		$Q(x_0)$	$\wedge$E 5
7		$\forall x\,Q(x)$	$\forall$I 5–6
8		$\forall x\,P(x) \wedge \forall x\,Q(x)$	$\wedge$I 4, 7

At lines 4 and 7, x_0 appears neither in the premise nor in the conclusion of that line, and no box is open. The second flagged box reuses the name x_0. That is allowed: the side condition at line 7 asks only about the premise, the open assumptions, and $\forall x\,Q(x)$, and the earlier box, with its lines 2 and 3, closed at line 3 and is not available at line 7. (A fresh y_0 would do equally well.)

($\dashv$)

1		$\forall x\,P(x) \wedge \forall x\,Q(x)$	
2		$\forall x\,P(x)$	$\wedge$E 1
3		$\forall x\,Q(x)$	$\wedge$E 1
4	x_0	$P(x_0)$	$\forall$E 2
5		$Q(x_0)$	$\forall$E 3
6		$P(x_0) \wedge Q(x_0)$	$\wedge$I 4, 5
7		$\forall x\,(P(x) \wedge Q(x))$	$\forall$I 4–6

At line 7, x_0 is in neither the premise nor the conclusion, and no box is open. The two derivations prove the two sequents, as claimed.

■

The first derivation is the answer to Section 5.3's stall: its lines 1–4 derive $\forall x\,P(x)$ from $\forall x\,(P(x) \wedge Q(x))$. The proposition says that $\forall$ distributes over $\wedge$.

Fine print: flagged boxes in a derivation. Definition 1.4.2 (Section 1.4) said that a box is a block of consecutive lines whose first line is an assumption. We now extend that definition, and two things change. First, a box may be **flagged**: it carries a name, a variable written at its top, and a flagged box need **not** open with an assumption. The box of $\forall$I has none; the box of the rule for using "there exists" in Section 5.6 is flagged and has one. Second, the rules that close a flagged box carry side conditions about its name, such as the one above. Everything else in Definition 1.4.2 stays word for word: two boxes are disjoint or one is inside the other, a box ends at one of its lines, and the scope rule decides which lines and boxes are available where. A flagged box with no assumption adds nothing to what is assumed; it only marks where

the name is in use. When a flagged box immediately contains another box, as in the swap of two quantifiers in the examples above, the range starts at the first line of that inner box. The right-hand derivation of "Why each clause", derivation (c) above and Exercise B1 below use the convention.

5.4.4 Every element of the empty set

Section 5.1 read $\forall x \in \emptyset\ \varphi$ as true whatever φ says, because there is no element to test, and promised a derivation. Section 2.1 met the same phenomenon for implications: $\neg P \vdash P \to Q$ (Proposition 2.1.2, the derived rule written VI). With $\forall$I the two meet.

Let A be a set with no elements. What that means is the premise $\forall x\, \neg(x \in A)$: no object is in A. From it, "every element of A is greater than 5", which by Definition 5.1.1 is $\forall x\, (x \in A \to x > 5)$, has a four-line derivation:

1	$\forall x\, \neg(x \in A)$	
	x_0 ⋮	
3	$x_0 \in A \to x_0 > 5$	
4	$\forall x\, (x \in A \to x > 5)$	

1	$\forall x\, \neg(x \in A)$	
2	x_0 $\neg(x_0 \in A)$	$\forall$E 1
	⋮	
3	$x_0 \in A \to x_0 > 5$	
4	$\forall x\, (x \in A \to x > 5)$	

Here is how the search grows it, and the result:

1	$\forall x\, \neg(x \in A)$	
2	x_0 $\neg(x_0 \in A)$	$\forall$E 1
3	$x_0 \in A \to x_0 > 5$	VI 2
4	$\forall x\, (x \in A \to x > 5)$	$\forall$I 2–3

Line 3 uses VI (Proposition 2.1.2) as one line, with the formulas $x_0 \in A$ and $x_0 > 5$ for P and Q; every derived rule of Chapter 2 may be used in this way on first-order formulas. Nothing in the derivation used $x > 5$: with any formula in place of $x > 5$ in which x_0 does not appear, the same four lines go through, so "every element of an empty set has property φ" is derivable for every φ, as Section 5.1 said.

Second reading · *Why a derived rule may stand in a flagged box*

Using VI inside the box is allowed because the derivation of Proposition 2.1.2 uses only ¬E, ⊥E and →I, whose schemas apply to any formulas (Definition 1.2.1), and it has no flagged box, so pasting it into the box flagged x_0 gives a derivation in the rules themselves. From now on, any derived rule of Chapter 2 is used in this way on first-order formulas.

One more flagged box gives a statement about upper bounds:

1	$\forall x\, \neg(x \in A)$		
2	u_0 | x_0 | $\neg(x_0 \in A)$		$\forall$E 1
3	| | $x_0 \in A \to x_0 \leq u_0$		VI 2
4	| $\forall x\, (x \in A \to x \leq u_0)$		$\forall$I 2–3
5	$\forall u\, \forall x\, (x \in A \to x \leq u)$		$\forall$I 2–4

Over $\mathbb{R}$, line 5 says that every real number u is an upper bound of A: every element of A is at most u. For an empty set this holds for every u, because there are no elements to check. At line 4, x_0 appears in neither the premise nor $\forall x\, (x \in A \to x \leq u_0)$, and the only open box (flagged u_0) has no assumption. The name u_0 does appear in that conclusion, which is allowed, since the condition at line 4 is about x_0 only. At line 5, u_0 appears in neither the premise nor $\forall u\, \forall x\, (x \in A \to x \leq u)$, and no box is open.

5.4.5 Common errors

Three errors with $\forall$I are common enough to name now.

- **Checking instances.** Verifying φ at $x = 0, 1, 2$, or at a thousand values, proves $\varphi[0/x]$, $\varphi[1/x]$, $\varphi[2/x]$, and so on, and not $\forall x\, \varphi$. The rule needs one proof about a fresh name, not many proofs about particular terms. Exercise C2 shows how plausible this error can look.
- **Generalizing over a name in an open assumption.** Inside "suppose $x_0 > 0$", the name x_0 is not arbitrary. The left-hand derivation of the pair in "Why each clause" shows where this leads.
- **Generalizing over a name in a premise.** A premise about x_0, such as "x_0 is even", makes x_0 special. The first derivation of "Why each clause" shows it.

Warning. "Let x be arbitrary" is a promise about every line that follows, and the most common way to break it is an assumption made on the way. "Let $n \in \mathbb{N}$ and suppose n is even. Then n^2 is even. Since n was arbitrary, every n^2 is even" is false at $n = 1$. What the argument proves is "for every n, **if** n is even, then n^2 is even": the $\to$I box on "n is even" must close **before** the $\forall$I step, so that the assumption is no longer open.

Quick check. In the derivation below, line 6 is justified by $\forall$I. Is it legal? If not, which clause of the side condition fails, and can you find a reading in which the premises are true and line 6 is false?

1	$R(c, y_0)$		
2	$\forall x\, \forall y\, (R(x, y) \to R(y, x))$		
3	y_0 | $\forall y\, (R(c, y) \to R(y, c))$		$\forall$E 2
4	| $R(c, y_0) \to R(y_0, c)$		$\forall$E 3
5	| $R(y_0, c)$		$\to$E 4, 1
6	$\forall y\, R(y, c)$		$\forall$I 3–5

Solution. No. Lines 1–5 are legal; in particular line 5 may cite line 1, which lies in no box. But at line 6 the flag y_0 appears in the premise on line 1, so the first clause fails: line 5 was proved about y_0 using a fact about y_0, and y_0 was not arbitrary. A reading: over $\mathbb{Z}$, let $R(x, y)$ be $x \neq y$, which is symmetric, so premise 2 is true; let c name 0 and y_0 name 1, so premise 1, $0 \neq 1$, is true. Line 6 says that every integer differs from 0, which is false at $y = 0$.

Example. A Proof to Check

Here is a proof. Is its claim true? Which step fails?

> *Claim.* Some real number is at least as large as every real number.
>
> *Proof.* Let x be a real number, and put $M = x + 1$. Then $x \leq M$. Since x was arbitrary, $x \leq M$ for every real x, so M is such a number.

Solution. The claim is false: whatever M is proposed, $M + 1$ is a larger real number. The faulty step is "since x was arbitrary". As a derivation, the box is flagged x_0 and ends in $x_0 \leq x_0 + 1$, a line of arithmetic; the proof then writes $\forall x\,(x \leq M)$, and M stands for $x_0 + 1$:

1	x_0 |	$x_0 \leq x_0 + 1$	arith
2		$\forall x\,(x \leq x_0 + 1)$	$\forall$I 1–1

Line 2 breaks the third clause: the flag is still in the conclusion. The box earns $\forall x\,(x \leq x+1)$, with **every** occurrence of the flag replaced: true, and about no single M. The name M was chosen after x, so it may depend on x, which is the dependence Section 5.1 read off the order of quantifiers. With Section 5.5's rule, $\exists$I inside the box and then $\forall$I give "every real number has one at least as large", with the quantifiers in the safe order.

5.4.6 The rule card

Example. Instances of $\forall$I

In each row the box's flag is in no premise and no open assumption.

flag	*the box's last line*	*the new line*
x_0	$P(f(x_0)) \vee \neg R(x_0, c)$	$\forall x\,(P(f(x)) \vee \neg R(x, c))$
ε_0	$\varepsilon_0 > 0 \to \exists\delta\,(\delta > 0 \wedge \delta < \varepsilon_0)$	$\forall\varepsilon\,(\varepsilon > 0 \to \exists\delta\,(\delta > 0 \wedge \delta < \varepsilon))$
y_0	$\forall x\,(R(x, y_0) \to R(x, g(y_0)))$	$\forall y\,\forall x\,(R(x, y) \to R(x, g(y)))$

Not an instance. A box flagged x_0 ends in $R(x_0, x_0)$. The line $\forall x\,R(x, x_0)$ is not written by $\forall$I: it keeps the flag in its conclusion (the third clause). The rule puts x in place of **every** occurrence of the flag: $\forall x\,R(x, x)$.

Chapters 1 and 2's thirteen rules are on the card at the end of Section 3.1; here are this chapter's so far.

Rule	Shape	Side condition	You write	Lean
∀E	$\forall x\,\varphi \Longrightarrow \varphi[t/x]$	t free for x in φ	"In particular, for $x = t$, . . . "; "applying the hypothesis to t, . . . "	`h t` (or `specialize h t`)
∀I	box $[x_0]$ . . . $\varphi[x_0/x] \Longrightarrow \forall x\,\varphi$	x_0 in no premise, in no open assumption, not in $\forall x\,\varphi$	"Let x be arbitrary. . . . Since x was arbitrary, . . . "	`intro x`

Other forms of the opening phrase are "let $x \in \mathbb{R}$", "fix $\varepsilon > 0$", "take any $n \in \mathbb{N}$"; the closing phrase is often left out, since the reader knows that a proof which began by fixing x proves a statement about every x. In Lean, `intro x` turns the goal `forall x, P x` into the goal `P x` with a new name `x` in context, about which nothing is assumed: Lean's version of opening the flagged box. The same tactic opened the →I box in Section 1.4, and the two boxes are, to Lean, one kind of thing.

In the Givens/Goal table, ∀I puts a fresh name among the Givens and makes $\varphi[x_0/x]$ the goal.

Warning (Not allowed: ∀I). Each step below is wrong; the tag names the clause of the side condition of Definition 5.4.1 that it breaks.

- Generalize over a flag that a premise mentions: the first clause.
- Generalize over a flag that an open assumption mentions: the second clause (the pair in "Why each clause").
- Write a conclusion that still mentions the flag: the third clause (Example (A Proof to Check)).
- Cite a line inside a flagged box that has closed: the scope rule of Definition 1.4.2, as for every box.

5.4.7 The search method: one more step

- **A goal $\forall x\,\varphi$.** Open a flagged box with a fresh name x_0, one that appears in no premise, in no open assumption, and not in the goal. The new goal, the box's last line, is $\varphi[x_0/x]$.

This step is **safe**, in the sense of Section 3.1. If $\forall x\,\varphi$ can be derived at all, then so can $\varphi[x_0/x]$, by one more ∀E at x_0; so opening the box never closes off a derivation that would otherwise exist. Take the step as soon as the goal is universal, and choose a flag that appears nowhere yet, so that the side condition cannot fail.

Remark (Looking ahead). Chapter 8 collects the errors of this section, with others, in a catalogue of steps a proof may not take.

5.4.8 Exercises

5.4.8.1 A. Check your understanding

Exercise. A1

Three derivations, D1, D2 and D3, are used in part (b).

D1:

1	$\forall x\,(P(x) \to Q(x))$	
2	$P(x_0)$	
3	x_0 | $P(x_0) \to Q(x_0)$	$\forall$E 1
4	| $Q(x_0)$	$\to$E 3, 2
5	$\forall x\, Q(x)$	$\forall$I 3–4

D2:

1	$\forall x\, \forall y\, R(x,y)$	
2	$\forall y\, R(x_0,y)$	$\forall$E 1
3	x_0 | $R(x_0,x_0)$	$\forall$E 2
4	$\forall x\, R(x,x)$	$\forall$I 3–3

D3:

1	| $R(x_0,c)$	asm, for $\to$I
2	| x_0 | $R(x_0,c)$	R 1
3	| $\forall x\, R(x,c)$	$\forall$I 2–2
4	$R(x_0,c) \to \forall x\, R(x,c)$	$\to$I 1–3

(a) Write the schema of $\forall$I with its side condition.

(b) For each of D1, D2, D3, decide whether its $\forall$I line is legal. If not, name the clause of the side condition that fails.

(c) True or false, with a reason: "fresh" means that x_0 appears nowhere in the derivation outside its flagged box.

(d) In the derivation of the ($\vdash$) half of Proposition 5.4.1, the name x_0 flags two boxes. Why does the side condition allow this?

Solution.

(a) A box flagged x_0, with no assumption, from line i to line j, whose last line j is $\varphi[x_0/x]$, a line of the flagged box itself; then line k: $\forall x\,\varphi$, justified "$\forall$I i–j". At line k the name x_0 appears in no premise, in no assumption of a box still open at line k, and not in $\forall x\,\varphi$ (Definition 5.4.1).

(b) *D1:* not legal. At line 5 the flag x_0 appears in the premise on line 2 (the first clause). Over $\mathbb{Z}$, with $P(x)$ and $Q(x)$ both read as "x is even" and x_0 naming 2, both premises

are true and line 5, "every integer is even", is false.
D2: legal. Line 2 mentions x_0 outside the box, and that is allowed: it is neither a premise nor an assumption, and the conclusion $\forall x\, R(x, x)$ does not contain x_0. Line 3 cites line 2, which is available, since no box contains it. (It is also correct in meaning: line 2 came from a premise that says nothing about x_0, so it holds whatever x_0 names.)
D3: not legal. At line 3 the assumption $R(x_0, c)$ on line 1 is still open, and x_0 appears in it (the second clause). Line 4, which closes the box legally, then says "if x_0 is below c, everything is below c": false over $\mathbb{Z}$ with R read as $<$, c as 5 and x_0 as 1.

(c) False. The side condition asks about the premises, the assumptions of boxes still open at the ∀I line, and the conclusion of that line. Other lines may mention x_0, as line 2 of D2 does, and a closed flagged box may share its name with a later one.

(d) At line 7, which closes the second box, the side condition asks whether x_0 appears in the premise (no), in an assumption of an open box (no box is open), and in $\forall x\, Q(x)$ (no). The first box's lines mention x_0, but they are neither premises nor open assumptions, and the box closed at line 3, so its lines are not even available at line 7.

Exercise. A2: Drill

Each part takes a line or two. The parts run from single steps, through a missing line, a justification column and a judgement on one step, to whole derivations.

(a) What does opening a flagged box say, in the words of a proof?

(b) A box flagged x_0 ends in $P(x_0) \wedge Q(x_0)$, and x_0 is in no premise and no open assumption. Which line may ∀I write?

(c) Line 9 is $\forall x\, (P(x) \vee Q(f(x)))$, justified "∀I 4–8", the box flagged u_0. What is line 8?

(d) Justify each line after the premise: 1 $\forall x\, (P(x) \wedge Q(x))$, premise; a box flagged x_0 holding 2 $P(x_0) \wedge Q(x_0)$, 3 $Q(x_0)$, 4 $P(x_0)$, 5 $Q(x_0) \wedge P(x_0)$; then 6 $\forall x\, (Q(x) \wedge P(x))$.

(e) The premise is $R(x_0)$. May a box flagged x_0 that ends in $R(x_0)$ be closed with $\forall x\, R(x)$? If not, which clause does it break?

(f) May a flag be the constant c?

(g) A box flagged y_0 ends in $y_0 \leq y_0$. Is $\forall x\, (x \leq y_0)$ a legal ∀I line? If not, which clause does it break?

(h) Inside the box of an assumption $Q(x_0)$, a box flagged x_0 ends in $Q(x_0)$. May ∀I close it? If not, which clause does it break?

(i) Derive $\forall x\, P(x) \vdash \forall y\, P(y)$.

(j) Derive $\forall x\, (P(x) \wedge Q(x)) \vdash \forall x\, Q(x)$.

(k) Derive $\forall x\, P(x),\ \forall x\, Q(x) \vdash \forall x\, (Q(x) \wedge P(x))$.

(l) Derive $\vdash \forall x\, (P(x) \to P(x))$.

Solution.

(a) "Let x_0 be arbitrary."

(b) $\forall x\, (P(x) \wedge Q(x))$.

(c) $P(u_0) \vee Q(f(u_0))$, the instance at the flag.

(d) 2 ∀E 1 (at x_0); 3 ∧E 2; 4 ∧E 2; 5 ∧I 3, 4; 6 ∀I 2–5.

(e) No, the first clause: x_0 appears in a premise.

(f) No: a flag is always a variable.

(g) No, the third clause: y_0 appears in the conclusion.

(h) No, the second clause: x_0 appears in an open assumption.

(i) A box flagged y_0: 2 $P(y_0)$, ∀E 1; then 3 $\forall y\, P(y)$, ∀I 2–2.

(j) A box flagged x_0: 2 $P(x_0) \wedge Q(x_0)$, ∀E 1; 3 $Q(x_0)$, ∧E 2; then 4 $\forall x\, Q(x)$, ∀I 2–3.

(k) A box flagged x_0: 3 $Q(x_0)$, ∀E 2; 4 $P(x_0)$, ∀E 1; 5 $Q(x_0) \wedge P(x_0)$, ∧I 3, 4; then 6 $\forall x\,(Q(x) \wedge P(x))$, ∀I 3–5.

(l) A box flagged x_0, and inside it the box of 1 $P(x_0)$; 2 $P(x_0) \to P(x_0)$, →I 1–1; then 3 $\forall x\,(P(x) \to P(x))$, ∀I 1–2 (cited from line 1, the convention for a flagged box that immediately contains another).

5.4.8.2 B. Practice

Exercise. B1: A chain, for every x

A search for $\forall x\,(P(x) \to Q(x)),\ \forall x\,(Q(x) \to R(x)) \vdash \forall x\,(P(x) \to R(x))$ has reached the stage below. *Name the move, then complete it, and check the side condition at line 9.*

1	$\forall x\,(P(x) \to Q(x))$	
2	$\forall x\,(Q(x) \to R(x))$	
3	x_0 | $P(x_0)$	asm, for →I
	⋮	
8	$P(x_0) \to R(x_0)$	?
9	$\forall x\,(P(x) \to R(x))$	?

Solution. The moves are *let x be arbitrary* and then the →I box; inside, instantiate both premises at x_0 and fire twice.

The goal is universal, so open a box flagged x_0; the new goal $P(x_0) \to R(x_0)$ is an implication, so open a →I box on $P(x_0)$; the new goal is $R(x_0)$. Now Section 5.3's exercise B1 runs inside the boxes, with x_0 for a.

1	$\forall x\,(P(x) \to Q(x))$		
2	$\forall x\,(Q(x) \to R(x))$		
3	x_0 | $P(x_0)$		asm, for $\to$I
4		$P(x_0) \to Q(x_0)$	$\forall$E 1
5		$Q(x_0)$	$\to$E 4, 3
6		$Q(x_0) \to R(x_0)$	$\forall$E 2
7		$R(x_0)$	$\to$E 6, 5
8		$P(x_0) \to R(x_0)$	$\to$I 3–7
9	$\forall x\,(P(x) \to R(x))$		$\forall$I 3–8

The flagged box immediately contains the $\to$I box, so line 9 cites it from line 3. At line 9: x_0 appears in neither premise; the assumption $P(x_0)$ mentions x_0, but its box closed at line 7, so it is not open at line 9; and x_0 does not appear in $\forall x\,(P(x) \to R(x))$. The order of the boxes matters: the $\to$I box is inside the flagged box and closes first.

Exercise. B2: Contraposition under a quantifier, used later

(a) Give a derivation of $\forall x\,(P(x) \to Q(x)) \vdash \forall x\,\neg Q(x) \to \forall x\,\neg P(x)$.

(b) Try to derive the converse, $\forall x\,\neg Q(x) \to \forall x\,\neg P(x) \vdash \forall x\,(P(x) \to Q(x))$. Say where the search stalls, and which side condition stops it.

Solution.

(a) The goal is an implication: a $\to$I box on $\forall x\,\neg Q(x)$, with goal $\forall x\,\neg P(x)$. That is universal: a box flagged x_0, with goal $\neg P(x_0)$. That is a negation: a $\neg$I box on $P(x_0)$, with goal $\bot$. Inside, instantiate both universal lines at x_0 and look for a clash: $P(x_0)$ and the instance of the premise give $Q(x_0)$, and the assumption $\forall x\,\neg Q(x)$ at x_0 gives $\neg Q(x_0)$.

1	$\forall x\,(P(x) \to Q(x))$		
2		$\forall x\,\neg Q(x)$	asm, for $\to$I
3		x_0 | $P(x_0)$	asm, for $\neg$I
4			$P(x_0) \to Q(x_0)$	$\forall$E 1
5			$Q(x_0)$	$\to$E 4, 3
6			$\neg Q(x_0)$	$\forall$E 2
7			$\bot$	$\neg$E 5, 6
8			$\neg P(x_0)$	$\neg$I 3–7
9			$\forall x\,\neg P(x)$	$\forall$I 3–8
10	$\forall x\,\neg Q(x) \to \forall x\,\neg P(x)$			$\to$I 2–9

At line 9, the open box is the $\to$I box of line 2, whose assumption $\forall x\,\neg Q(x)$ does

not contain x_0; the premise and the conclusion do not contain it either. So the side condition holds, even though the flagged box sits inside an open box.

(b) The goal is universal, so open a box flagged x_0, then a $\to$I box on $P(x_0)$, with goal $Q(x_0)$. The only premise is an implication, $\forall x\, \neg Q(x) \to \forall x\, \neg P(x)$. Using it needs its hypothesis $\forall x\, \neg Q(x)$, which nothing supplies; and its conclusion, $\forall x\, \neg P(x)$, would clash with $P(x_0)$ and close the case by $\bot$E. So everything turns on deriving $\forall x\, \neg Q(x)$. The goal $Q(x_0)$ is atomic and nothing produces it, so the method's last resort is an RAA box on $\neg Q(x_0)$. Inside, the available facts about Q are $\neg Q(x_0)$, and to reach $\forall x\, \neg Q(x)$ from it we would close a flagged box over x_0. The side condition forbids it: x_0 appears in the open assumptions $P(x_0)$ and $\neg Q(x_0)$. A fresh name y_0 does not help either, since nothing tells us anything about $\neg Q(y_0)$. The search stalls at the second clause of the side condition.

A stalled search proves nothing by itself; a cleverer derivation might exist. Section 7.2 settles the question with a structure in which the premise is true and the conclusion false, and shows that the converse does not follow.

Exercise. B3: A number is at most its absolute value, for every number

Example (A number is at most its absolute value) (Section 1.6) derived $|x| > x \vee |x| = x$ for one fixed real x from three premises about that x: $x \geq 0 \vee x < 0$, $x \geq 0 \to |x| = x$, and $x < 0 \to |x| > x$. Turn it into a derivation of

$$\forall x\, (|x| > x \vee |x| = x)$$

from premises that are universal statements. Explain why the premises of Example (A number is at most its absolute value), copied with x_0 in place of x, will not do.

Solution. The goal is universal, so the derivation of Section 1.6 must run inside a box flagged x_0. Its three premises mention the number it is about. So here the premises are the three facts **for every** x, and inside the box $\forall$E at x_0 gives each fact about x_0.

1		$\forall x\,(x \geq 0 \lor x < 0)$	
2		$\forall x\,(x \geq 0 \to \lvert x\rvert = x)$	
3		$\forall x\,(x < 0 \to \lvert x\rvert > x)$	
4	x_0	$x_0 \geq 0 \lor x_0 < 0$	$\forall$E 1
5		$x_0 \geq 0 \to \lvert x_0\rvert = x_0$	$\forall$E 2
6		$x_0 < 0 \to \lvert x_0\rvert > x_0$	$\forall$E 3
7		$x_0 \geq 0$	asm, for $\lor$E
8		$\lvert x_0\rvert = x_0$	$\to$E 5, 7
9		$\lvert x_0\rvert > x_0 \lor \lvert x_0\rvert = x_0$	$\lor$I 8
10		$x_0 < 0$	asm, for $\lor$E
11		$\lvert x_0\rvert > x_0$	$\to$E 6, 10
12		$\lvert x_0\rvert > x_0 \lor \lvert x_0\rvert = x_0$	$\lor$I 11
13		$\lvert x_0\rvert > x_0 \lor \lvert x_0\rvert = x_0$	$\lor$E 4, 7–9, 10–12
14		$\forall x\,(\lvert x\rvert > x \lor \lvert x\rvert = x)$	$\forall$I 4–13

Lines 7–13 are the derivation of Example (A number is at most its absolute value) with x_0 for x, its premises replaced by lines 4–6. At line 14, x_0 appears in none of the premises, no box is open (the two case boxes closed at lines 9 and 12), and x_0 is not in the conclusion. The derivation has 14 lines.

If instead the premises were $x_0 \geq 0 \lor x_0 < 0$, $x_0 \geq 0 \to \lvert x_0\rvert = x_0$ and $x_0 < 0 \to \lvert x_0\rvert > x_0$, the box could reach line 13 just as well, but line 14 would break the first clause: x_0 would appear in a premise. And rightly so. Premises about x_0 are facts about one number, and a derivation from them proves something about that number only. The universal premises are true facts about every real number, and that is what entitles the conclusion to speak about every real number.

Exercise. B4: Back to Exercise C2 of Section 5.3

(a) Derive $\forall x\,P(x) \vdash \forall x\,P(x)$ with $\forall$E and $\forall$I, not R. Why can the flagged box not be flagged x?

(b) Explain why no flag at all gives a derivation of $P(x) \vdash \forall x\,P(x)$ by this route, and give a counter-reading to the sequent.

Solution.

	1		$\forall x\,P(x)$	
(a)	2	x_0	$P(x_0)$	$\forall$E 1
	3		$\forall x\,P(x)$	$\forall$I 2–2

With the flag x, the conclusion $\forall x\,P(x)$ contains x (the letter after the quantifier

counts: that is what "appear" means), so the third clause fails, whatever else holds.

(b) With the flag x, the line $P(x)$ is available, but x appears in the premise and in the conclusion. With any other flag x_0, the box needs the line $P(x_0)$, and the search finds nothing that gives it: the premise is about x, and no line mentions x_0. A stalled search proves nothing by itself, since a cleverer derivation might exist; the counter-reading settles it. Over $\mathbb{Z}$, read P as "even" and let x name 2: the premise is true and $\forall x\, P(x)$ is false. (Section 7.2 makes this a counter-model.)

Marking scheme. (10 points) - 3: the derivation of (a). - 3: the flag x fails "not in the conclusion", with the reason. - 2: (b) both cases of the flag. - 2: the counter-reading, with the value of x. Fatal errors (0 for the step): claiming the flag x is fine because x is bound in the conclusion; treating $P(x)$ and $P(x_0)$ as the same line.

Exercise. B5: Derive or break: properties of a relation, used later

A two-place predicate R is **reflexive** if $\forall x\, R(x,x)$, **irreflexive** if $\forall x\, \neg R(x,x)$, **symmetric** if $\forall x\, \forall y\, (R(x,y) \to R(y,x))$, **asymmetric** if $\forall x\, \forall y\, (R(x,y) \to \neg R(y,x))$, and **transitive** if $\forall x\, \forall y\, \forall z\, ((R(x,y) \land R(y,z)) \to R(x,z))$. Read each claim with these formulas as its premises and conclusion, and derive it, or break it with a counter-reading. Then say which two claims are converses, and which claim repairs which.

(a) asymmetric $\vdash$ irreflexive.

(b) irreflexive $\vdash$ asymmetric.

(c) irreflexive, transitive $\vdash$ asymmetric.

(d) symmetric, transitive $\vdash$ reflexive.

(e) symmetric, transitive, $\forall x\, R(x, f(x)) \vdash$ reflexive.

(f) reflexive $\vdash$ symmetric.

Solution.

(a) Derive. The goal is universal, so a flagged box; its goal $\neg R(x_0, x_0)$ is a negation, so a $\neg$I box; the premise at x_0, twice, gives the clash.

1	$\forall x\, \forall y\, (R(x,y) \to \neg R(y,x))$		
2	x_0	$R(x_0, x_0)$	asm, for $\neg$I
3		$\forall y\, (R(x_0,y) \to \neg R(y,x_0))$	$\forall$E 1
4		$R(x_0,x_0) \to \neg R(x_0,x_0)$	$\forall$E 3
5		$\neg R(x_0,x_0)$	$\to$E 4, 2
6		$\bot$	$\neg$E 2, 5
7		$\neg R(x_0,x_0)$	$\neg$I 2–6
8	$\forall x\, \neg R(x,x)$		$\forall$I 2–7

At line 8, x_0 is not in the premise, no box is open, and x_0 is not in $\forall x\, \neg R(x,x)$.

(b) Break. The domain $\{0,1\}$, R true of $(0,1)$ and $(1,0)$ only. No element is related to itself; $R(0,1)$ and $R(1,0)$ both hold.

(c) Derive: two flags, then an $\to$I box, then a $\neg$I box. Transitivity at x_0, y_0, x_0 turns the two assumptions into $R(x_0, x_0)$, which irreflexivity refutes.

1	$\forall x\, \neg R(x, x)$	
2	$\forall x\, \forall y\, \forall z\, ((R(x, y) \land R(y, z)) \to R(x, z))$	
3	x_0 | y_0 | $R(x_0, y_0)$	asm, for $\to$I
4	$R(y_0, x_0)$	asm, for $\neg$I
5	$\forall y\, \forall z\, ((R(x_0, y) \land R(y, z)) \to R(x_0, z))$	$\forall$E 2
6	$\forall z\, ((R(x_0, y_0) \land R(y_0, z)) \to R(x_0, z))$	$\forall$E 5
7	$(R(x_0, y_0) \land R(y_0, x_0)) \to R(x_0, x_0)$	$\forall$E 6
8	$R(x_0, y_0) \land R(y_0, x_0)$	$\land$I 3, 4
9	$R(x_0, x_0)$	$\to$E 7, 8
10	$\neg R(x_0, x_0)$	$\forall$E 1
11	$\bot$	$\neg$E 9, 10
12	$\neg R(y_0, x_0)$	$\neg$I 4–11
13	$R(x_0, y_0) \to \neg R(y_0, x_0)$	$\to$I 3–12
14	$\forall y\, (R(x_0, y) \to \neg R(y, x_0))$	$\forall$I 3–13
15	$\forall x\, \forall y\, (R(x, y) \to \neg R(y, x))$	$\forall$I 3–14

At line 14, y_0 is in no premise, the only open box (flagged x_0) has no assumption, and y_0 is not in the conclusion; at line 15 the same holds for x_0, with no box open.

(d) Break. The domain $\{0, 1\}$, R true of $(0, 0)$ only. R is symmetric and transitive (the only pair is $(0, 0)$), and $R(1, 1)$ fails.

(e) Derive. Inside the box, the partner $f(x_0)$ plays the part that b played in Exercise C1 of Section 5.3.

1		$\forall x\,\forall y\,(R(x,y) \to R(y,x))$	
2		$\forall x\,\forall y\,\forall z\,((R(x,y) \land R(y,z)) \to R(x,z))$	
3		$\forall x\,R(x, f(x))$	
4	x_0	$R(x_0, f(x_0))$	$\forall$E 3
5		$\forall y\,(R(x_0,y) \to R(y,x_0))$	$\forall$E 1
6		$R(x_0, f(x_0)) \to R(f(x_0), x_0)$	$\forall$E 5
7		$R(f(x_0), x_0)$	$\to$E 6, 4
8		$\forall y\,\forall z\,((R(x_0,y) \land R(y,z)) \to R(x_0,z))$	$\forall$E 2
9		$\forall z\,((R(x_0, f(x_0)) \land R(f(x_0),z)) \to R(x_0,z))$	$\forall$E 8
10		$(R(x_0, f(x_0)) \land R(f(x_0), x_0)) \to R(x_0,x_0)$	$\forall$E 9
11		$R(x_0, f(x_0)) \land R(f(x_0), x_0)$	$\land$I 4, 7
12		$R(x_0,x_0)$	$\to$E 10, 11
13		$\forall x\,R(x,x)$	$\forall$I 4–12

At line 13, x_0 is in no premise, no box is open, and x_0 is not in the conclusion.

(f) Break. The domain $\{0,1\}$, R true of $(0,0)$, $(1,1)$ and $(0,1)$. R is reflexive; $R(0,1)$ holds and $R(1,0)$ does not.

(g) and (b) are converses, and only (a) derives. (d) is the trap "symmetric and transitive give reflexive", which breaks at an element related to nothing, and (e) repairs it by giving every element a partner $f(x)$ (Exercise C1 of Section 5.3 shows the same for one element that has a partner).

Marking scheme. (10 points) - 2 each for (a), (c), (e): flags opened in the order read off the goal, every $\forall$E at the flag (or at $f(x_0)$) the box needs, the side condition checked once at the last $\forall$I. - 1 each for (b), (d), (f): a counter-reading with the domain and every pair of R listed, each premise and the conclusion checked. - 1: the closing sentence. Fatal errors (0 for the claim): a $\forall$I over a flag that an open assumption mentions; a counter-reading that is not symmetric (in (d)) or not reflexive (in (f)).

5.4.8.3 C. Going deeper

Exercise. C1: How fresh is fresh?

(a) Give a derivation of $\forall x\,\forall y\,R(x,y) \vdash R(x_0,x_0) \land \forall x\,R(x,x)$ in which the line $R(x_0,x_0)$ is written **before** a flagged box that uses the same name x_0. Check every clause of the side condition at the $\forall$I line.

(b) A stronger condition is sometimes proposed: "x_0 occurs nowhere in the derivation outside its box". Does your derivation meet it? Explain why the conclusion of your derivation is nonetheless correct, so that the stronger condition forbids more than it needs to.

(c) Could the last line of the flagged box in (a) be obtained by R, repeating the earlier line $R(x_0,x_0)$? Is the result a derivation?

Solution (Hint 1). The goal is a conjunction. Derive the left conjunct first, by two ∀E lines outside any box.

Solution (Hint 2). For the right conjunct, open a box flagged x_0 and instantiate the premise at x_0 again inside it. At the ∀I line, list the premises, the open assumptions and the conclusion, and look for x_0 in each.

Solution.

	1	$\forall x\,\forall y\,R(x,y)$	
	2	$\forall y\,R(x_0,y)$	∀E 1
	3	$R(x_0,x_0)$	∀E 2
(a)	4	x_0 | $\forall y\,R(x_0,y)$	∀E 1
	5	| $R(x_0,x_0)$	∀E 4
	6	$\forall x\,R(x,x)$	∀I 4–5
	7	$R(x_0,x_0)\wedge\forall x\,R(x,x)$	∧I 3, 6

The side condition at line 6. *Premises:* the only premise is line 1, $\forall x\,\forall y\,R(x,y)$, and x_0 does not appear in it. *Open assumptions:* no box is open at line 6, so there are none. *Conclusion:* $\forall x\,R(x,x)$ does not contain x_0. The last line of the box, line 5, is $R(x,x)[x_0/x]$ and is a line of the box itself. So line 6 is legal, and the derivation has 7 lines. Lines 2, 3 and 7 mention x_0 outside the box, and the side condition does not ask about them.

(b) No: lines 2, 3 and 7 mention x_0 outside its box. Yet the conclusion is correct. The reason the side condition looks only at premises and open assumptions is that these are the only lines taken without proof, so they are the only lines that can single out what x_0 names; every other available line is derived from them. Lines 2 and 3 were derived from a premise that does not mention x_0; they hold whatever x_0 names. So the argument in the box really is about an arbitrary object, and $\forall x\,R(x,x)$ follows. The stronger condition would reject this derivation, although renaming the box's flag to a new letter z_0 (lines 4–5 becoming $\forall y\,R(z_0,y)$ and $R(z_0,z_0)$) gives a derivation that meets it and has the same last line. A condition that rejects a correct derivation for the spelling of a name forbids more than safety needs.

(c) Yes. R may copy into the box any line available there, and line 3 lies in no box, so it is available at every later line. With line 5 justified "R 3" (and line 4 deleted), the box is a single line $R(x_0,x_0)$, and the side condition at the ∀I line is checked exactly as in (a): nothing in it looks at where the box's last line came from. The derivation is legal, and it is correct for the reason in (b): line 3 was derived from a premise that says nothing about x_0.

Marking scheme. (10 points) - 2: the left conjunct by two ∀E lines at x_0, outside any box. - 2: a flagged box with ∀E at the flag and ∀I citing it. - 2: the side condition checked clause by clause: premise, open assumptions, conclusion. - 2: (b) explains that lines outside the box that do not rest on premises or assumptions about x_0 cannot make x_0 special. Alternatively, 2 for showing that renaming the box's flag to a new letter gives a derivation that meets the

stronger condition and has the same last line. - 2: (c) correct answer with the availability check for the repeated line. Fatal errors (0 for the step): claiming that line 3 makes line 6 illegal; checking the side condition against lines that are not premises, open assumptions or the conclusion.

Exercise. C2: Find the error

A student offers this proof.

> **Claim.** For every real number x, $x^2 - 2x + 1 \geq 0$.
>
> **Proof.** (1) Let x be a real number. (2) At $x = 0$: $0 - 0 + 1 = 1 \geq 0$. (3) At $x = 1$: $1 - 2 + 1 = 0 \geq 0$. (4) At $x = 2$: $4 - 4 + 1 = 1 \geq 0$. (5) So $x^2 - 2x + 1 \geq 0$, and since x was arbitrary, the claim holds for every real x.

(a) Is the claim true?

(b) Does the argument prove it? Name the step that fails and the rule it misuses.

(c) Show that the step can lead to a false conclusion: give a polynomial g for which the same checks at $x = 0, 1, 2$ succeed and $g(x) \geq 0$ fails for some real x.

(d) Point to a correct proof.

Solution (Hint 1). Write lines (2)–(4) as formulas. Which terms do they mention?

Solution (Hint 2). Write the student's proof as a derivation: open a box flagged x_0 for line (1), and write lines (2)–(4) as lines inside it. What would line (5) have to cite?

Solution.

(a) Yes. For every real x, $x^2 - 2x + 1 = (x - 1)^2 \geq 0$, since a square is never negative.

(b) No. Lines (2)–(4) are true, and line (1) opens a flagged box, "let x be arbitrary". The error is in line (5). With φ the formula $x^2 - 2 \cdot x + 1 \geq 0$, lines (2)–(4) establish $\varphi[0/x]$, $\varphi[1/x]$ and $\varphi[2/x]$: statements about the constants 0, 1, 2. None of them is $\varphi[x_0/x]$ for the flag. Inside the box, nothing about the arbitrary x was ever proved. The first sentence of (5), "so $x^2 - 2x + 1 \geq 0$", has the right form: it is $\varphi[x_0/x]$, with the student's x as the flag. But nothing the student wrote yields it: lines (2)–(4) are the only lines it rests on, and no rule takes $\varphi[0/x]$, $\varphi[1/x]$ and $\varphi[2/x]$ to $\varphi[x_0/x]$. So the closing "since x was arbitrary" applies ∀I to a box whose last line was never derived: lines (2)–(4) are all instances at constants, not the instance at the flag. Line (1) was idle: the argument never used the x it introduced.

(c) Take $g(x) = x(x - 1)(x - 2)$. The same checks succeed: $g(0) = 0 \geq 0$, $g(1) = 0 \geq 0$, $g(2) = 0 \geq 0$. But $g(-1) = (-1)(-2)(-3) = -6 < 0$, so "$g(x) \geq 0$ for every real x" is false. So the pattern of (5), concluding "for every x" from checks at a few constants, can lead from true checks to a false conclusion. It is not an instance of ∀I, as (b) showed. (A legal step never leads from true premises to a false conclusion; Section 7.2 states this.) The claim in (a) is true for a different reason, one that the argument never gave.

(d) The derivation under "A first flagged box" in this section: inside the box flagged x_0, $\forall$E applies "squares are not negative" at $x_0 - 1$, an arith line rewrites it as $x_0^2 - 2x_0 + 1 \geq 0$, and that line **is** $\varphi[x_0/x]$, so $\forall$I applies.

Marking scheme. (10 points) - 2: (a) the claim is true, with a reason (it is a square). - 3: (b) names step (5) and $\forall$I; says that lines (2)–(4) are instances at constants, not the instance at the flag. - 3: (c) a polynomial that passes the three checks and fails somewhere, with both computations shown. - 2: (d) a correct proof, generalizing over a statement proved about the arbitrary x. Fatal errors (0 for the step): answering (b) with "the claim is false"; a "counterexample" in (c) that fails one of the three checks or is non-negative everywhere.

5.5 Proving "There Exists": Exhibit a Witness

Sections 5.3 and 5.4 gave the two rules for "for every": $\forall$E uses a universal statement, and $\forall$I proves one. This section starts on the other quantifier. It gives the rule that **proves** an existential statement, $\exists$I, and the habit that goes with it: when the goal is "there exists", do not guess the object, but wait until a line of the derivation shows it. The sequent $\forall x\, P(x) \vdash \exists x\, P(x)$, which no rule so far can finish, is finished here in three lines.

5.5.1 Showing that something exists

How does a mathematician prove that there is a natural number N with $1/(N+1) < 1/100$? By naming one. "Take $N = 100$: then $1/(N+1) = 1/101$, and $1/101 < 1/100$." Nobody asks for more, and nobody objects that $N = 1000$ would also have done. The proof exhibits a **witness**, checks that it has the property, and stops.

Look at what the proof does with the witness at the end. The statement proved, "there is an N with $1/(N+1) < 1/100$", does not say which N; the 100 still in it belongs to the bound $1/100$, not to N. The number was needed to establish the claim and is then forgotten: a reader of the conclusion learns that some N works, not which one.

Chapter 1 used the same pattern from its first page. It recalled that an integer n is **even** when $n = 2k$ for **some** integer k, and said that the word "some" has rules of its own, to come in Chapter 5. Here is the first of them. To show that 6 is even, take $k = 3$: then $6 = 2 \cdot 3$. The statement "6 is even" is $\exists k\, (6 = 2 \cdot k)$, and the proof passes from the instance $6 = 2 \cdot 3$ to it. The second rule for "some", how to **use** "$n = 2k$ for some k", is the subject of Section 5.6.

There is a helpful way to see the step, which Section 5.1's reading of $\exists$ suggests. Over a domain with three elements named a, b, c, the sentence $\exists x\, P(x)$ says what the disjunction

$$P(a) \vee P(b) \vee P(c)$$

says (grouped to the right, as in Section 4.5): at least one of the three has P. An existential is a long "or", with one disjunct for each object of the domain. Proving a disjunction from one of its sides is $\vee$I (Definition 1.5.1), and proving an existential from one instance is the rule of this section. The analogy is only a guide (a domain such as $\mathbb{N}$ has infinitely many elements, and a formula cannot be infinitely long), but it predicts the rule's behavior exactly: it throws information away, and it must never be run backwards.

5.5.2 The rule of existential introduction

To prove that something has a property, show that one particular thing has it.

Definition 5.5.1. Existential Introduction

i	$\varphi[t/x]$	
k	$\exists x\,\varphi$	$\exists\mathrm{I}\ i$

The rule given by this schema is called $\exists$I (**existential introduction**). Here φ is **any** formula, x any variable, and t **any** term, subject to one side condition: t **is free for** x **in** φ (Definition 5.2.6). Line i must be available at line k (Definition 1.4.2).

- **Schema:** the block above: one cited line, no box, no fresh name.
- **Reading:** from a line that says something about the term t, write "something has that property": mark some occurrences of t with a variable x to get φ, and write $\exists x\,\varphi$.
- **Example:** from $6 = 2 \cdot 3$, marking the 3 gives $\exists k\,(6 = 2 \cdot k)$.

Which occurrences? Any you like, as the table below shows; the term t is the **witness**, and it may be a constant, a compound term or a variable: a variable used elsewhere in the derivation, a flag such as x_0, or even the variable x itself. Check two things. First, $\varphi[t/x]$ must be **exactly** the cited line: choose a variable x that does not appear in the line and mark some occurrences of t with it, and this marking recipe guarantees it; with an x that does appear, check directly (from $R(x, c)$, marking c gives $R(x, x)$, and $R(x, x)[c/x]$ is $R(c, c)$, not the line). Second, t must be free for x in φ, the condition of Definition 5.2.6 that $\forall$E also carries. With a bound, $\exists x \in A\ \varphi$, the witness needs the line $t \in A$ too; "A bounded witness" below shows the two steps.

One line, several readings. Let line i be $R(c, c)$, with R a two-place predicate symbol and c a constant. The witness is c, and c occurs twice. Each occurrence may be marked with x or left alone, and each choice is a different φ:

φ	$\varphi[c/x]$ (the cited line)	the new line $\exists x\,\varphi$
$R(x, c)$	$R(c, c)$	$\exists x\,R(x, c)$
$R(c, x)$	$R(c, c)$	$\exists x\,R(c, x)$
$R(x, x)$	$R(c, c)$	$\exists x\,R(x, x)$

All three are legal steps. Read over $\mathbb{N}$ with $R(x, y)$ as "$x \leq y$" and c as 5: the cited line says $5 \leq 5$, and the three useful conclusions say that something is at most 5, that 5 is at most something, and that something is at most itself. Each is a true and weaker statement.

Second reading · *A witness that is not marked*

A fourth row is legal too:

φ	$\varphi[c/x]$ (the cited line)	the new line $\exists x\,\varphi$
$R(c, c)$	$R(c, c)$	$\exists x\,R(c, c)$

The last row is the **degenerate** choice: φ has no free x, so $\varphi[c/x]$ is φ (Section 5.2) and the new line $\exists x\, R(c,c)$ quantifies over nothing. It is legal, and it says nothing that line i did not.

Why this rule. Under the reading of Section 5.1, $\exists x\, \varphi$ is true when at least one element of the domain, put in for x, makes φ true. If $\varphi[t/x]$ is true, then the object that t names is such an element, so $\exists x\, \varphi$ is true. The side condition is what makes "$\varphi[t/x]$ says about t what φ says about x" correct: that was the content of "free for" in Section 5.2.

$\exists$I *forgets the witness, and keeps the truth.*

Quick check. Line 1 is $x_0 + 1 > x_0$. Give an $\exists$I line with the witness x_0, and one with the witness $x_0 + 1$.

Solution. Marking both occurrences of x_0 with y gives $\varphi = (y + 1 > y)$ and the line $\exists y\,(y + 1 > y)$; marking the term $x_0 + 1$ gives $\varphi = (y > x_0)$ and the line $\exists y\,(y > x_0)$. In both, y appears nowhere in line 1, and the formula has no quantifier, so the witness is free for y.

A nonexample: capture. Let line 1 be $\forall y\, R(y,y)$. With $\varphi = \forall y\, R(x,y)$ and $t = y$, the definition of substitution gives $\varphi[y/x] = \forall y\, R(y,y)$, which is line 1. So it is tempting to write $\exists x\, \forall y\, R(x,y)$ by $\exists$I.

1	$\forall y\, R(y,y)$	
2	$\exists x\, \forall y\, R(x,y)$	$\exists$I 1

Line 2 breaks the side condition: the free occurrence of x in φ lies in the scope of $\forall y$, and y is the variable of the term t, so y is **not** free for x in φ. The step would be wrong. Read over $\mathbb{Z}$ with $R(x,y)$ as "$x \leq y$": line 1 says that every integer is at most itself, which is true, and line 2 says that some integer is at most every integer, that is, that there is a least integer, which is false. The y put in for x was meant to name one object, and it was captured by $\forall y$ and made to range over all of them.

Warning. The witness must be **written on a line**. "There is an N with $1/(N+1) < 1/100$: take N large enough" is not an application of $\exists$I, because "large enough" is not a term, and the rule asks for an instance $\varphi[t/x]$ on an available line. In a written proof, "large enough" is a promise to name the witness, and the promise is kept by a term such as $N = 100$, followed by the check that it works ($1/101 < 1/100$). Until then nothing has been proved.

Quick check. Let line 1 be $2 + 2 = 4$. Which of

$$\exists x\,(x + x = 4), \qquad \exists x\,(x + 2 = 4), \qquad \exists x\,(2 + 2 = x)$$

may line 2 be, justified by one application of $\exists$I to line 1? For each that may, name φ and the witness t.

Solution. All three. For $\exists x\,(x + x = 4)$, take $\varphi = (x + x = 4)$ and $t = 2$: both 2's of line 1 are marked. For $\exists x\,(x + 2 = 4)$, take $\varphi = (x + 2 = 4)$ and $t = 2$: only the first 2 is marked, and the second is written out in φ. For $\exists x\,(2 + 2 = x)$, take $\varphi = (2 + 2 = x)$ and $t = 4$: the witness is a different term. In each case $\varphi[t/x]$ is exactly line 1, and φ has no quantifier, so t is free for x.

5.5.3 From "for every" to "there exists"

The sequent $\forall x\,P(x) \vdash \exists x\,P(x)$ looks like the easiest in the chapter, and it has a small puzzle in it. To use $\exists$I, the goal needs an instance $P(t)$ on a line. The premise gives an instance for every term, by $\forall$E. But the sequent mentions no constant, so which term?

Proposition 5.5.1. Every Implies Some

$\forall x\,P(x) \vdash \exists x\,P(x)$.

Proof. The following is a derivation.

1	$\forall x\,P(x)$	
2	$P(x)$	$\forall$E 1
3	$\exists x\,P(x)$	$\exists$I 2

Line 2 is $\forall$E at the term x: with $\varphi = P(x)$, the instance $\varphi[x/x]$ is $P(x)$, and the variable x is free for x in every formula (Section 5.2). Line 3 is $\exists$I with the same φ and the witness x. This proves the proposition. ∎

The puzzle is answered by the paragraph "Which occurrences?" after Definition 5.5.1: the witness may be a variable, the variable x itself included, so no constant is needed. The line $P(x)$ has x free; it says that P holds of whatever x names. Nothing in the derivation says what that is, and nothing needs to. The same three lines prove $\forall x\,\varphi \vdash \exists x\,\varphi$ for every formula φ, since $\varphi[x/x] = \varphi$ (Section 5.2).

Remark. The rules accept this derivation in any case; it leads from a true premise to a true conclusion only because every domain has **at least one** element, the convention of Section 5.1. Over an empty domain, $\forall x\,P(x)$ would be true with nothing to check and $\exists x\,P(x)$ false with no witness; the convention enters at line 2, where $P(x)$ treats x as naming an object, and there must be one to name.

↻ **Second reading** · *The empty set, again*

This does not clash with the empty set of Section 5.1, where $\forall x \in \emptyset\ \varphi$ is true and $\exists x \in \emptyset\ \varphi$ is false. A bounded quantifier is an abbreviation (Definition 5.1.1): $\forall x \in \emptyset\ \varphi$ is $\forall x\,(x \in \emptyset \to \varphi)$,

and $\exists x \in \emptyset\ \varphi$ is $\exists x\,(x \in \emptyset \wedge \varphi)$. The domain is still non-empty; it is the set $\emptyset$ that has no elements. Running the three lines on the bounded premise gives $\forall$E at x, the line $x \in \emptyset \to \varphi$, and then $\exists$I gives only $\exists x\,(x \in \emptyset \to \varphi)$, an existential with $\to$ inside. No line $x \in \emptyset$ is available to fire the implication, so the bounded "every implies some" is not derived, and it should not be.

5.5.4 An existential goal: postpone the witness

The search method of Section 3.1 asks two questions at each gap: what is the goal, and what are the resources? Sections 5.3 and 5.4 added a step for each side of $\forall$. Now $\exists$ can be a goal.

If the goal is $\exists x\,\varphi$, its last line can be $\exists$I, and then the new goal is $\varphi[t/x]$ for a term t **of our choosing**. That is the same situation as a goal $\varphi \vee \psi$, where $\vee$I left us to choose a side, and the danger is the same. A wrong t can leave a goal that does not follow from the resources, although the existential does. Section 3.1 met the danger for $\vee$ and answered it by postponing; the answer here is the same.

Postpone the witness: choose the term t of a goal $\exists x\,\varphi$ only when an available line shows an instance $\varphi[t/x]$.

Quick check. Premises: $\forall x\,(P(x) \to Q(x+1))$ and $P(2)$. The goal is $\exists y\,Q(y)$. Which witness will the search find, and after which steps?

Solution. The term $2+1$: $\forall$E at 2 gives $P(2) \to Q(2+1)$, $\to$E gives $Q(2+1)$, and only then does $\exists$I have an instance to cite. (Writing the witness as 3 instead needs the line $2+1=3$ and the equality rule of Section 5.7.)

Until then, work on the premises and on the other goals, and watch what terms appear. Often the witness is a term that only appears late, built by $\forall$E or $\to$E from the premises. Here are four derivations, in order of difficulty, each with an Idea that says why each rule was chosen.

Example. Derivations with existential introduction

Show that:

(a) $P(c) \wedge Q(c) \vdash \exists x\,(P(x) \wedge Q(x))$;

(b) $R(c,c) \vdash \exists x\,R(x,c) \wedge \exists x\,R(x,x)$; and decide whether one $\exists$I step leads from $\exists x\,R(x,x)$ to $\exists x\,R(x,c)$;

(c) $\forall x\,R(x,x) \vdash \forall x\,\exists y\,R(x,y)$;

(d) $\forall x\,(P(x) \to Q(f(x))),\ P(c) \vdash \exists y\,Q(y)$.

Solution.

(a) *Idea.* The goal is an existential, and the premise already is an instance of a suitable φ: mark both occurrences of c. One goal-directed step finishes.

1	$P(c) \wedge Q(c)$	
2	$\exists x\,(P(x) \wedge Q(x))$	$\exists$I 1

Here $\varphi = P(x) \wedge Q(x)$ and $t = c$. Marking only the first c, with $\varphi = P(x) \wedge Q(c)$, would give the legal but weaker line $\exists x\,(P(x) \wedge Q(c))$, which is not the goal.

(b) *Idea.* The goal is a conjunction, so it splits into two goals (Section 1.2), and each is an existential with the line $R(c, c)$ as an instance, read in two different ways (the first and third rows of the table above). The goal at the bottom (left); then the two subgoals that $\wedge$I asks for (right).

1	$R(c, c)$	
	⋮	
4	$\exists x\, R(x, c) \wedge \exists x\, R(x, x)$	?

1	$R(c, c)$	
	⋮	
2	$\exists x\, R(x, c)$	?
	⋮	
3	$\exists x\, R(x, x)$	?
4	$\exists x\, R(x, c) \wedge \exists x\, R(x, x)$	$\wedge$I 2, 3

1	$R(c, c)$	
2	$\exists x\, R(x, c)$	$\exists$I 1
3	$\exists x\, R(x, x)$	$\exists$I 1
4	$\exists x\, R(x, c) \wedge \exists x\, R(x, x)$	$\wedge$I 2, 3

Line 2 uses $\varphi = R(x, c)$ and line 3 uses $\varphi = R(x, x)$, both with the witness c.
Now the second question. From the line $\exists x\, R(x, x)$, the step to $\exists x\, R(x, c)$ is **not** an $\exists$I step. The new line is $\exists x\, \varphi$ with $\varphi = R(x, c)$, so the cited line would have to be $R(t, c)$ for some term t. The line $\exists x\, R(x, x)$ is not of that form: its main symbol is $\exists x$, not R. The rule passes from an **instance** to an existential, never from one existential to another. Nor is the step safe by any other route: read over $\mathbb{Z}$ with $R(x, y)$ as "$x \cdot y = 1$" and c as 0, the line $\exists x\, R(x, x)$ is true (take $x = 1$), and $\exists x\, R(x, c)$, "some integer times 0 is 1", is false.

(c) *Idea.* The goal is a $\forall$, so the step of Section 5.4 opens a flagged box with a fresh name x_0, and the new goal is $\exists y\, R(x_0, y)$. That is an existential: postpone it and look at the premise. The premise is a $\forall$; instantiate it at the term the goal mentions, x_0. The result, $R(x_0, x_0)$, shows the witness: it is x_0 itself, a variable used as a term, as x was in Proposition 5.5.1. The box with its postponed goal (left); then the instance that shows the witness (right).

1	$\forall x\, R(x, x)$		
	x_0 ⋮		
3	$\exists y\, R(x_0, y)$	?	
4	$\forall x\, \exists y\, R(x, y)$	?	

1	$\forall x\, R(x, x)$		
2	x_0 $R(x_0, x_0)$	$\forall$E 1	
3	$\exists y\, R(x_0, y)$	?	
4	$\forall x\, \exists y\, R(x, y)$	?	

1		$\forall x\, R(x,x)$	
2	x_0	$R(x_0,x_0)$	$\forall$E 1
3		$\exists y\, R(x_0,y)$	$\exists$I 2
4		$\forall x\, \exists y\, R(x,y)$	$\forall$I 2–3

Line 3 uses $\varphi = R(x_0,y)$, the variable y, and the witness x_0; it marks the second occurrence of x_0 and keeps the first. Line 4 is $\forall$I with the flag x_0 (Definition 5.4.1): x_0 appears in no premise, there is no open assumption, and x_0 does not appear in line 4. The $\exists$I inside the box needed no box of its own, although its witness is a fresh name. Read over $\mathbb{N}$ with $R(x,y)$ as "$x \leq y$": every number is at most itself, so every number is at most something.

(d) *Idea.* The goal $\exists y\, Q(y)$ is an existential, and no line yet has the form $Q(t)$. Postpone, and work forwards. The premise $\forall x\,(\dots)$ is a $\forall$; the only term in sight is c, so instantiate at c (Section 5.3), and $\to$E with the premise $P(c)$ fires. The result $Q(f(c))$ shows the witness: the compound term $f(c)$. In the Givens/Goal table:

Givens	Goal
$\forall x\,(P(x) \to Q(f(x)))$, $P(c)$	$\exists y\, Q(y)$: postponed, no instance yet
$+\ P(c) \to Q(f(c))$ ($\forall$E at c)	$\exists y\, Q(y)$
$+\ Q(f(c))$ ($\to$E)	$\exists y\, Q(y)$ ✓ by $\exists$I, witness $f(c)$

The postponed goal (left); then the two forward steps, after which the witness is visible (right).

1	$\forall x\,(P(x) \to Q(f(x)))$	
2	$P(c)$	
	⋮	
5	$\exists y\,Q(y)$	?

1	$\forall x\,(P(x) \to Q(f(x)))$	
2	$P(c)$	
3	$P(c) \to Q(f(c))$	$\forall$E 1
4	$Q(f(c))$	$\to$E 3, 2
5	$\exists y\,Q(y)$	?

1	$\forall x\,(P(x) \to Q(f(x)))$	
2	$P(c)$	
3	$P(c) \to Q(f(c))$	$\forall$E 1
4	$Q(f(c))$	$\to$E 3, 2
5	$\exists y\,Q(y)$	$\exists$I 4

Line 5 uses $\varphi = Q(y)$ and $t = f(c)$. Now see what choosing the witness first would have done. The term in sight at the start is c, and "take $y = c$" leaves the goal $Q(c)$. Nothing gives it: the premise speaks of Q only at terms of the form $f(\dots)$, and $Q(c)$ need not hold at all. Read over $\mathbb{Z}$ with P as "odd", Q as "even", $f(x)$ as $x + 1$, and c as 3: the premises say that the successor of every odd integer is even and that 3 is odd, the conclusion says that some integer is even, and the witness the derivation finds is $f(3) = 4$, while $Q(c)$ says that 3 is even. The search stalls exactly because it chose the witness before the derivation had produced it.

The four Ideas share a pattern. In (a) and (b) the witness stood on a premise line from the start; in (c) and (d) it appeared only after other rules had run, and in (d) it was a term that no line mentioned at the start. Postponing costs nothing in (a) to (c), where the right witness was already a term in sight when the existential goal appeared (in (c) the goal names x_0), and it is necessary in (d), where the only term in sight at the start, c, is the wrong witness.

5.5.5 How this rule looks in a mathematical proof

In a written proof, $\exists$I is the word "take", followed by a check. "6 is even: take $k = 3$; then $6 = 2 \cdot 3$." Over $\mathbb{Z}$, with one line of school arithmetic justified "arith" as in Section 5.3:

1	$6 = 2 \cdot 3$	arith
2	$\exists k\,(6 = 2 \cdot k)$	$\exists$I 1

Line 2 marks the 3 with k, and keeps the 6 and the 2: $\varphi = (6 = 2 \cdot k)$. In the same way "7 is odd: take $k = 3$, since $7 = 2 \cdot 3 + 1$" is an $\exists$I with $\varphi = (7 = 2 \cdot k + 1)$.

A bounded witness. Analysis states its existentials with bounds. "There is an $N \in \mathbb{N}$ with $1/(N + 1) < 1/100$" abbreviates $\exists N\,(N \in \mathbb{N} \wedge 1/(N + 1) < 1/100)$ (Definition 5.1.1), so the witness must satisfy both conjuncts. Taking $N = 100$ is two steps: $\wedge$I puts the bound and the inequality together, and $\exists$I forgets the witness.

1	$100 \in \mathbb{N}$	arith
2	$1/(100+1) < 1/100$	arith
3	$100 \in \mathbb{N} \wedge 1/(100+1) < 1/100$	$\wedge$I 1, 2
4	$\exists N\,(N \in \mathbb{N} \wedge 1/(N+1) < 1/100)$	$\exists$I 3

Line 4 marks two of the three occurrences of 100 in line 3 and keeps the one on the right. The witness 100 is the least that works, since $N = 99$ gives $1/100$, which is not less than $1/100$; but the proof does not need the least witness, only one.

A witness built from what is available. Often the witness is a term made from other names. To show that for every integer n there is an odd integer greater than $2n$, a proof says "let n be an integer; $2n+1$ will do, since it is odd and greater than $2n$". The words "let n be an integer" open the flagged box of Section 5.4, and "$2n+1$ will do" is $\exists$I with the compound witness $2n+1$, as in (d) above. In general, the witness may be any term the reader can check.

Warning. $\exists$I runs one way only. From $P(c)$ it gives $\exists x\,P(x)$; from $\exists x\,P(x)$ it gives nothing, and in particular not $P(c)$. "Some integer is even" is true and "3 is even" is false: the existential does not say **which** object has the property, so it cannot be read back as a statement about a named one. Using an existential takes a rule of its own, with a box, and that is Section 5.6.

5.5.6 The rule card

Before the card, the rule at work on lines larger than $R(c,c)$. The schema sees the cited line as $\varphi[t/x]$ however much is inside t and φ.

Example. Instances of $\exists$I

Cited line	φ	Witness t	New line
$P(f(f(c))) \wedge Q(f(c))$	$P(f(x)) \wedge Q(x)$	$f(c)$	$\exists x\,(P(f(x)) \wedge Q(x))$
$\lvert a-b\rvert < \varepsilon/2$	$\lvert a-b\rvert < \delta$	$\varepsilon/2$	$\exists \delta\,(\lvert a-b\rvert < \delta)$
$(n+1)\cdot(n+1) =$ $n\cdot n+2\cdot n+1$	$m\cdot m = n\cdot n+2\cdot n+1$	$n+1$	$\exists m\,(m\cdot m =$ $n\cdot n+2\cdot n+1)$

In each row $\varphi[t/x]$ is exactly the cited line, and the marked places are whole occurrences of t (in the third, both factors on the left; the n's on the right are not occurrences of $n+1$, and they stay).

Not an instance. From $P(c) \wedge Q(d)$, the line $\exists x\,(P(x) \wedge Q(x))$ is not written by $\exists$I: $\varphi[t/x]$ puts **one** term in every marked place, and no single t turns $P(x) \wedge Q(x)$ into $P(c) \wedge Q(d)$.

Quick check. Line 1 is $Q(g(c), c)$. Write the $\exists$I line with the witness c that marks only the c inside $g(c)$, and name its φ.

Solution. $\exists x\, Q(g(x), c)$, with $\varphi = Q(g(x), c)$; $\varphi[c/x]$ is line 1.

Chapters 1 and 2's thirteen rules are on the card at the end of Section 3.1; here are this chapter's so far. Nothing later depends on the last column, and Section 5.7 collects the whole table.

Rule	Shape	Side condition	You write	Lean
∀E	$\forall x\, \varphi \Longrightarrow \varphi[t/x]$	t free for x in φ	"In particular, for $x = t$, ..."; "applying the hypothesis to t, ..."	`h t` (or `specialize h t`)
∀I	box $[x_0]$ ... $\varphi[x_0/x] \Longrightarrow \forall x\, \varphi$	x_0 in no premise, in no open assumption, not in $\forall x\, \varphi$	"Let x be arbitrary. ... Since x was arbitrary, ..."	`intro x`
∃I	$\varphi[t/x] \Longrightarrow \exists x\, \varphi$	t free for x in φ	"Take $x = t$. Then ..."; "$2k + 1$ will do."	`use t` (or `Exists.intro t h`)

The phrase names the witness and then checks it. A proof that names the witness without the check ("take $N = 100$.") has written the line $\exists x\, \varphi$ without the instance $\varphi[t/x]$ it cites; the reader must supply the instance, and should be able to.

In the Givens/Goal table, ∃I ticks off the goal $\exists x\, \varphi$ once a Given is an instance $\varphi[t/x]$; until then the goal row waits.

Warning (Not allowed: ∃I). Each step is offered as one application of ∃I; each breaks the rule where shown.

- "From $P(c) \wedge Q(d)$, write $\exists x\, (P(x) \wedge Q(x))$": the cited line must be exactly $\varphi[t/x]$, with one term t in every marked place.
- "From $\forall y\, (y < y + 1)$, write $\exists x\, \forall y\, (y < x)$": the side condition; the witness $y + 1$ is not free for x in $\forall y\, (y < x)$.
- "From $P(t)$, write $\exists x \in A\ P(x)$": the cited line must be the instance; the bounded form is $\exists x\, (x \in A \wedge P(x))$, whose instance is $t \in A \wedge P(t)$ ("A bounded witness").

The search method: one more step. Each rule section adds a step to the method of Section 3.1.

- **Goal $\exists x\, \varphi$.** Postpone. Work on the premises and on other goals until an available line is an instance $\varphi[t/x]$, with t free for x in φ; then ∃I. The witness may be a compound term that appears only after other rules have run, as $f(c)$ did in (d).

5.5.7 Exercises

5.5.7.1 A. Check your understanding

Exercise. A1

(a) State ∃I as a schema, with its side condition.

(b) Line 1 of a derivation is $R(c, f(c))$. Which of the following may line 2 be, by one application of ∃I to line 1? For each that may, give φ and the witness t; for each that may not, say why.
(i) $\exists x\, R(x, f(c))$; (ii) $\exists x\, R(c, x)$; (iii) $\exists x\, R(x, f(x))$; (iv) $\exists x\, R(x, x)$; (v) $\exists x\, R(f(c), x)$.

(c) Line 1 of a derivation is $\forall y\, (y \leq y \cdot y)$. Is $\exists x\, \forall y\, (y \leq x \cdot x)$ a legal ∃I step from line 1? Read both lines over $\mathbb{Z}$: which is true?

(d) True or false, with a reason: from the line $\exists x\, P(x)$, one application of ∃I gives $P(c)$.

(e) The current goal of a search is $\exists x\, \varphi$ and no available line is an instance of φ. What does the search method say to do, and why not choose a witness now?

Solution. (a) From $\varphi[t/x]$ on an available line i, write $\exists x\, \varphi$ on line k, justified "∃I i". Side condition: t is free for x in φ.

(b) (i) Yes: $\varphi = R(x, f(c))$, $t = c$; only the first c is marked. (ii) Yes: $\varphi = R(c, x)$, $t = f(c)$; the witness is the compound term. (iii) Yes: $\varphi = R(x, f(x))$, $t = c$; both occurrences of c are marked, and $\varphi[c/x]$ is $R(c, f(c))$. (iv) No: the conclusion fixes $\varphi = R(x, x)$, whose instances $R(t, t)$ have equal arguments, while the arguments c and $f(c)$ of line 1 differ. (v) No: the conclusion fixes $\varphi = R(f(c), x)$, whose instances $R(f(c), t)$ have $f(c)$ as first argument, while line 1 has c there. (Read over $\mathbb{N}$ with $R(x, y)$ as "$x < y$", $f(x)$ as $x + 1$, and c as 0: line 1 says $0 < 1$, and (v) says that something exceeds 1, true but not by this step; (iv) says that some number is less than itself, which is false.)

(c) No. With $\varphi = \forall y\, (y \leq x \cdot x)$, the formula $\varphi[y/x]$ is line 1, so the witness would be the variable y; but the free occurrences of x in φ lie in the scope of $\forall y$, so y is not free for x in φ. Over $\mathbb{Z}$, line 1 is true: for $y \leq 0$, $y \leq 0 \leq y \cdot y$, and for $y \geq 1$, $y \leq y \cdot y$ after multiplying $1 \leq y$ by $y > 0$. The other line is false: whatever integer x is proposed, $y = x \cdot x + 1$ gives $y > x \cdot x$. So the step would lead from a true line to a false one, which is what the side condition prevents.

(d) False. ∃I concludes an existential from an instance; it never concludes an instance. The line $P(c)$ does not follow from $\exists x\, P(x)$ at all: over $\mathbb{Z}$ with P as "even" and c as 3, the first is true and the second false.

(e) Postpone the goal: work on the premises (∀E at the terms the goal mentions, →E, and so on) until a line $\varphi[t/x]$ appears, then use ∃I. A witness chosen now leaves the goal $\varphi[t/x]$, which may not follow from the resources even when $\exists x\, \varphi$ does, as "take $y = c$" did in Example (Derivations with existential introduction) (d).

Exercise. A2: Drill

Each part takes a line or two. The parts run from single steps, through a missing line, a justification column and steps to judge, to short derivations.

(a) Line 1 is $3 + 4 = 7$. Name φ and t so that ∃I gives $\exists x\, (x + 4 = 7)$.

(b) Line 2 is $c + c = 2 \cdot c$. Write the ∃I lines with the witness c that mark (i) all three occurrences of c, (ii) the two on the left only.

(c) Line 1 is $P(f(c))$. Give two different $\exists$I conclusions.

(d) Line 5 is $\exists y\,(y > x_0+1)$, justified "$\exists$I 4". Give two possible lines 4, with their witnesses.

(e) Justify each line: 1 $\forall x\,(P(x) \to Q(f(x)))$, premise; 2 $P(c)$, premise; 3 $P(c) \to Q(f(c))$; 4 $Q(f(c))$; 5 $P(c) \wedge Q(f(c))$; 6 $\exists x\,(P(x) \wedge Q(f(x)))$.

(f) Line 1 is $R(c,c)$. Which of $\exists x\,R(x,c)$, $\exists x\,R(c,x)$, $\exists x\,R(x,y)$ can one $\exists$I **not** give?

(g) Line 1 is $\forall y\,(y = y)$. Is $\exists x\,\forall y\,(x = y)$ a legal $\exists$I line?

(h) True or false: $\exists$I may cite $\exists x\,P(x)$ and write $P(c)$.

(i) Premises $P(a)$, $Q(b)$; goal $\exists x\,(P(x) \wedge Q(x))$. Should a witness be chosen now?

(j) Derive $P(c) \vdash \exists x\,(P(x) \vee Q(x))$.

(k) Derive $\forall x\,P(x) \vdash \exists x\,P(f(x))$.

(l) Derive $\forall x\,\forall y\,R(x,y) \vdash \exists x\,\exists y\,R(y,x)$.

Solution.

(a) $\varphi = (x + 4 = 7)$, $t = 3$.

(b) (i) $\exists x\,(x + x = 2 \cdot x)$; (ii) $\exists x\,(x + x = 2 \cdot c)$.

(c) $\exists x\,P(x)$ (witness $f(c)$) and $\exists x\,P(f(x))$ (witness c).

(d) Any line $t > x_0+1$ with t free for y (every term is, since the formula has no quantifier): for example $x_0 + 2 > x_0 + 1$ (witness $x_0 + 2$) or $7 > x_0 + 1$ (witness 7).

(e) 3 $\forall$E 1 (at c); 4 $\to$E 3, 2; 5 $\wedge$I 2, 4; 6 $\exists$I 5 (witness c, marking the c of $P(c)$ and the one inside $f(c)$).

(f) $\exists x\,R(x,y)$: its instances $R(t,y)$ have y in the second place, and line 1 has c.

(g) No: the witness would be y, which $\forall y$ captures. Over $\mathbb{Z}$ line 1 is true and the other line, "some integer equals every integer", false.

(h) False: $\exists$I goes from an instance to an existential, never back.

(i) No: neither a nor b has both properties on a line. In fact a counter-reading exists: the domain $\{a, b\}$, with P true of a only and Q of b only.

(j) 2 $P(c) \vee Q(c)$, $\vee$I 1; 3 $\exists x\,(P(x) \vee Q(x))$, $\exists$I 2.

(k) 2 $P(f(x))$, $\forall$E 1 (at the term $f(x)$); 3 $\exists x\,P(f(x))$, $\exists$I 2 (with $\varphi = P(f(x))$ and the witness x).

(l) 2 $\forall y\,R(x,y)$, $\forall$E 1 (at x); 3 $R(x,x)$, $\forall$E 2; 4 $\exists y\,R(y,x)$, $\exists$I 3 (witness x, marking the first place); 5 $\exists x\,\exists y\,R(y,x)$, $\exists$I 4 (witness x).

5.5.7.2 B. Practice

Exercise. B1: Postpone, then choose

A search for $\forall x\,(P(x) \to Q(x)),\ P(a) \vdash \exists x\,(P(x) \wedge Q(x))$ has reached the stage below. *Name the move, then complete it, saying for each line whether the goal or a premise chose it.*

1	$\forall x\,(P(x) \to Q(x))$	
2	$P(a)$	
	$\vdots$	
6	$\exists x\,(P(x) \wedge Q(x))$	?

Solution. The move is *postpone the witness.*

Idea. The goal is an existential: postpone it. The premise $\forall x\,(\dots)$ is instantiated at the only term in sight, a, and $\to$E with $P(a)$ gives $Q(a)$; both steps are premise-directed. Now the witness a is visible, and the instance the goal needs is $P(a) \wedge Q(a)$, which $\wedge$I builds. The last step is the postponed $\exists$I.

1	$\forall x\,(P(x) \to Q(x))$	
2	$P(a)$	
3	$P(a) \to Q(a)$	$\forall$E 1
4	$Q(a)$	$\to$E 3, 2
5	$P(a) \wedge Q(a)$	$\wedge$I 2, 4
6	$\exists x\,(P(x) \wedge Q(x))$	$\exists$I 5

Line 6 uses $\varphi = P(x) \wedge Q(x)$ and the witness a.

Exercise. B2

Over $\mathbb{R}$, with the premise $\forall x\,(x + (-x) = 0)$, derive $\forall x\,\exists y\,(x + y = 0)$, the true sentence (i) of Example (Swapping two quantifiers). Say what the witness is, and what it depends on.

Solution. Idea. The goal is a $\forall$: open a flagged box with a fresh name x_0 (Section 5.4); the new goal is $\exists y\,(x_0 + y = 0)$. Postpone it, and instantiate the premise at the term the goal mentions, x_0. The line $x_0 + (-x_0) = 0$ shows the witness: $-x_0$.

1		$\forall x\,(x + (-x) = 0)$	
2	x_0	$x_0 + (-x_0) = 0$	$\forall$E 1
3		$\exists y\,(x_0 + y = 0)$	$\exists$I 2
4		$\forall x\,\exists y\,(x + y = 0)$	$\forall$I 2–3

Line 3 uses $\varphi = (x_0 + y = 0)$ and the witness $-x_0$, which contains no variable bound in φ. At line 4, x_0 appears in no premise, there is no open assumption, and x_0 does not appear in the conclusion, so $\forall$I applies (Definition 5.4.1).

The witness $-x_0$ is built from x_0, so it depends on x_0: a different number for each x. That is legal, because $\exists$I is applied **inside** the box, where x_0 is available as a term. It is the box picture of the game of Section 5.1: the $\exists y$ is played after the $\forall x$, and may use it.

Exercise. B3

Over $\mathbb{Z}$, prove "there is an integer n with $n^2 = n + 2$" as a derivation whose only line not justified by a rule is one line of arithmetic justified "arith"; write n^2 as $n \cdot n$. Then write the proof as one sentence.

Solution. The witness 2 works, since $2 \cdot 2 = 4 = 2 + 2$.

1	$2 \cdot 2 = 2 + 2$	arith
2	$\exists n\,(n \cdot n = n + 2)$	$\exists$I 1

Line 2 uses $\varphi = (n \cdot n = n + 2)$ with the witness 2. Line 1 contains four 2's, and φ marks three of them: both factors on the left and the first term of the sum on the right. The last 2 is the constant of the statement and stays. Marking all four would give $\exists n\,(n \cdot n = n + n)$, a different statement.

In one sentence: *Take $n = 2$; then $n^2 = 4 = n + 2$.*

The witness -1 works too, since $(-1) \cdot (-1) = 1 = -1 + 2$; a proof needs only one.

Exercise. B4: Derive or break

For each claim, derive it with the rules of Chapters 1–2, $\forall$E, $\forall$I and $\exists$I, or break it with a counter-reading (Section 5.3).

(a) $\forall x\, P(x) \vdash \exists x\,(P(x) \vee Q(x))$.
(b) $\exists x\, P(x) \vdash \forall x\, P(x)$.
(c) $R(c,d) \wedge R(d,c) \vdash \exists x\, R(x,x)$.
(d) $R(c,d) \wedge R(d,c) \vdash \exists x\, \exists y\,(R(x,y) \wedge R(y,x))$.
(e) $\forall x\, R(x, f(x)) \vdash \forall x\, \exists y\, R(x,y)$.
(f) $\forall x\, R(x, f(x)) \vdash \exists y\, \forall x\, R(x,y)$.
(g) $\vdash \exists x\,(P(x) \to P(x))$.

Solution.

(a) Derive. Instantiate at the variable x, weaken, and forget the witness.

1	$\forall x\, P(x)$	
2	$P(x)$	$\forall$E 1
3	$P(x) \vee Q(x)$	$\vee$I 2
4	$\exists x\,(P(x) \vee Q(x))$	$\exists$I 3

Line 4 uses $\varphi = P(x) \vee Q(x)$ and the witness x.

(b) Break. The domain $\{0, 1\}$, with P true of 0 only. The premise holds (at 0), and the conclusion fails at 1.

(c) Break. The domain $\{0, 1\}$, with c naming 0, d naming 1, and R true of the pairs $(0, 1)$ and $(1, 0)$ only. The premise holds, and no element is related to itself.

(d) Derive. Two witnesses, one for each place, forgotten one at a time.

1	$R(c,d) \wedge R(d,c)$	
2	$\exists y\,(R(c,y) \wedge R(y,c))$	$\exists$I 1
3	$\exists x\,\exists y\,(R(x,y) \wedge R(y,x))$	$\exists$I 2

Line 2 uses $\varphi = R(c,y) \wedge R(y,c)$ and the witness d; line 3 uses $\varphi = \exists y\,(R(x,y) \wedge R(y,x))$ and the witness c, which is free for x since it is a constant.

(e) Derive. Fix x_0 first; the witness $f(x_0)$ is found after the flag.

1	$\forall x\,R(x, f(x))$		
2	x_0	$R(x_0, f(x_0))$	$\forall$E 1
3		$\exists y\,R(x_0, y)$	$\exists$I 2
4	$\forall x\,\exists y\,R(x,y)$		$\forall$I 2–3

At line 4, x_0 appears in no premise, no box is open, and x_0 is not in line 4.

(f) Break. The domain $\{0,1\}$, with f the identity ($f(0) = 0$, $f(1) = 1$) and R true of $(0,0)$ and $(1,1)$ only. The premise holds, since each element is related to itself; and no y has both $R(0,y)$ and $R(1,y)$.

(g) Derive. A box with its assumption as its only line, then $\exists$I at the variable x.

1	$P(x)$	asm, for $\to$I
2	$P(x) \to P(x)$	$\to$I 1–1
3	$\exists x\,(P(x) \to P(x))$	$\exists$I 2

Line 3 has the witness x; the variable names an object, since domains are not empty (Section 5.1).

The pairs: (c) and (d) are look-alikes (one witness for both places, against two witnesses); (e) and (f) differ only in the order of $\forall$ and $\exists$, the swap of Section 5.1, and the witness $f(x_0)$ in (e) depends on the flag; (b) is the converse of Proposition 5.5.1. Worked examples of these types: Example (Derivations with existential introduction); Proposition 5.5.1.

Marking scheme. (10 points)

- 1 each: (a), (b), (c), (d), (f), (g), with the derivation or the counter-reading in full.
- 2: (e), with the flag opened first and the witness $f(x_0)$ found after it.
- 2: the closing pairing.

Fatal errors (0 for the claim): an $\exists$I whose witness is not free for the variable; a "break" in which the premise is false.

5.5.7.3 C. Going deeper

Exercise. C1: The conclusion clause of for-all-introduction, used later

The side condition of $\forall$I (Definition 5.4.1) has three clauses about the flag x_0: it appears in no premise, in no assumption of a box still open at the $\forall$I line, and not in the conclusion $\forall x\,\varphi$. Show that the third clause is needed: give a "derivation" that breaks **only** that clause, from a premise that is true over $\mathbb{N}$ to a sentence (Definition 5.2.4) that is false over $\mathbb{N}$. Check

that every other step is legal.

Solution (Hint 1). The conclusion of the illegal ∀I still mentions the flag, so it is not yet a sentence. Which rule of this section can replace a name on a line by a bound variable?

Solution (Hint 2). Start from the premise $\forall x\,(x \leq x)$, and open a flagged box with the name y_0.

Solution.

1	$\forall x\,(x \leq x)$	
2	y_0 | $y_0 \leq y_0$	∀E 1
3	$\forall x\,(x \leq y_0)$	∀I 2–2
4	$\exists y\,\forall x\,(x \leq y)$	∃I 3

The illegal line. Line 3 is ∀I with the flag y_0 and $\varphi = (x \leq y_0)$: the instance $\varphi[y_0/x]$ is $y_0 \leq y_0$, the last line of the box. The flag y_0 appears in no premise (line 1 has no y_0), and there is no open assumption at all. But y_0 appears in the conclusion $\forall x\,(x \leq y_0)$, so the third clause fails, and only it.

The other lines are legal. Line 2 is ∀E at the term y_0, which is free for x in $x \leq x$ since that formula has no quantifier. Line 4 is ∃I with $\varphi = \forall x\,(x \leq y)$ and the witness y_0: $\varphi[y_0/y]$ is line 3, and y_0 is free for y in φ, since the only quantifier of φ is on x.

Over $\mathbb{N}$. Line 1 says that every natural number is at most itself: true. Line 4 says that some natural number is at least every natural number, that is, that there is a largest natural number: false, since for any y, $x = y + 1$ is larger.

What the clause prevents. The box proved $y_0 \leq y_0$, a fact about y_0 in which the two sides are the same name. Line 3 generalized only one side and kept the other, so it claimed that **one** fixed y_0 is at least every x: a claim about y_0 that the box never proved. The flag is "let y_0 be arbitrary"; it must not survive into the statement that "since y_0 was arbitrary" produces. Line 4 then shows the damage, by turning the leaked name into a witness.

Marking scheme. (10 points) - 3: a derivation whose only illegal line is a ∀I whose flag appears in its conclusion. - 2: checks the other two clauses at that line (no premise, no open assumption contains the flag). - 2: checks the remaining lines (here ∀E and ∃I, with free-for). - 3: shows the premise true and the conclusion false over $\mathbb{N}$, with the witness of the falsity $(y + 1)$. Fatal errors (0 for the step): a second illegal line (for instance a capture in ∃I); a premise that is false over $\mathbb{N}$.

Exercise. C2: A witness is not a name, used later

Explain why the rules so far (Chapters 1–2 and ∀E, ∀I, ∃I) cannot derive $\exists x\,P(x) \vdash P(c)$.

Go through the rules that could produce the last line $P(c)$, and say what each would need. Then say, in one sentence, why no rule for using $\exists$ **should** give $P(c)$.

Solution. The argument below is informal: it explains why every search stalls, and it does not by itself prove that no derivation exists. Section 7.2 gives a model in which the premise is true and $P(c)$ is false. With the soundness theorem stated there (and proved in the optional Section 7.3), that model shows no derivation exists.

Which rules could write $P(c)$. The line $P(c)$ is atomic, so no introduction rule writes it: $\wedge$I, $\to$I, $\vee$I, $\neg$I, $\leftrightarrow$I, $\forall$I and $\exists$I all write a line whose main symbol is their own connective or quantifier. The remaining rules take $P(c)$ out of something larger: $\wedge$E needs a conjunction with $P(c)$ as a side; $\to$E needs an implication $\psi \to P(c)$ and ψ; $\leftrightarrow$E a biconditional with $P(c)$ as a side; $\vee$E a disjunction and two boxes each ending in $P(c)$; $\forall$E a universal $\forall x\,\psi$ with $\psi[c/x] = P(c)$; $\neg$E writes only $\bot$, never $P(c)$; $\bot$E the line $\bot$; RAA a box from $\neg P(c)$ to $\bot$; and R the line $P(c)$ itself.

The derived rules of Chapter 2 abbreviate derivations in these rules (Definition 2.3.1), so they add no new way to write $P(c)$.

What the premise can supply. The only premise is $\exists x\,P(x)$, whose main symbol is $\exists$. No rule so far looks inside an existential: each elimination rule needs a cited line that is $\bot$ (for $\bot$E) or whose main symbol is $\wedge$, $\to$, $\vee$, $\leftrightarrow$, $\neg$ or $\forall$, and $\exists$I produces existentials rather than using them. So the premise can enter a derivation only as a whole: copied, built into a larger formula, set against a line $\neg\exists x\,P(x)$ by $\neg$E, or cited as the side ψ beside a line $\psi \to \chi$ or $\psi \leftrightarrow \chi$ by $\to$E or $\leftrightarrow$E. None of these reaches the information inside it, and the lines the rules above need (a conjunction containing $P(c)$, an implication ending in it, a biconditional with it as a side, a disjunction for $\vee$E, a universal with it as an instance, a contradiction) have nowhere to come from. A search therefore stalls at once.

Why it should stall. The premise says that **something** has P and does not say what; the conclusion says that the object named c has P. Over $\mathbb{Z}$ with P as "even" and c as 3, the premise is true and the conclusion false. So no rule for using an existential may give $P(c)$, and the rule of Section 5.6 does not: it names the witness with a **fresh** variable, about which nothing else is known, instead of with c.

Marking scheme. (10 points) - 3: identifies that $P(c)$ is atomic, so only elimination rules, $\bot$E, RAA or R can write it, and names what each needs. - 3: observes that no rule so far uses a line whose main symbol is $\exists$, so the premise cannot supply what those rules need. - 2: says that this explains the stall but is not a proof of non-derivability (that needs a counter-model together with soundness, Sections 7.2–7.3). - 2: gives a reading with the premise true and $P(c)$ false, and the one-sentence reason (the premise does not say which object has P). Fatal errors (0 for the step): claiming that $\exists$I read backwards gives $P(c)$; claiming that a failed search proves the sequent underivable.

5.6 Using "There Exists": Name a Witness

Section 5.5 gave the rule that proves an existential statement: exhibit a witness, and forget it. This section gives the rule that **uses** one, $\exists$E. It is the last of the four quantifier rules, and the subtlest, because an existential says that a witness exists without saying which one it is. The rule handles this with a box and a fresh name, as $\forall$I did, but the name now stands for an object

that is **not** arbitrary, and the side condition keeps it from being treated as if it were. With the rule in hand, the section moves an $\exists$ past a $\forall$, and shows exactly where the reverse move breaks down.

5.6.1 Using a witness whose identity is unknown

Chapter 2 proved that if n^2 is even then n is even (Theorem 2.1.3). One of the facts it rested on, (E3), says that the square of an odd integer is odd, and it was shown by taking "an integer k with $n = 2k + 1$" and computing with it. Chapter 2 noted that naming the k of (E3) lay outside its derivation: it is a quantifier step, whose rule had to wait. Here is the simplest proof with the same step.

> *If n is even, then $n + 2$ is even.* Since n is even, choose an integer k with $n = 2k$. Then $n + 2 = 2k + 2 = 2(k + 1)$, so $n + 2$ is even.

The hypothesis is $\exists k\,(n = 2 \cdot k)$. The proof does not know which integer k is; it does not need to. It gives the unknown witness a name, uses the one fact it has about it, $n = 2k$, and reaches a conclusion, "$n + 2$ is even", in which the name no longer appears. The witness $k + 1$ of the last step is forgotten by $\exists$I, as in Section 5.5.

Why is this correct? Two tempting shortcuts show what the proof is careful **not** to do.

- It does not name the witness by a term that already means something. From "some integer is even" we may not conclude "3 is even": the existential does not say that the witness is 3, and, as the warning "$\exists$I runs one way only" of Section 5.5 noted, it cannot be read back as a statement about a named object.
- It does not let the conclusion mention the name. If it did, the conclusion would be a statement about k, and outside the argument k refers to nothing; the existential promised only that **some** k works.

So the name must be new, the argument may use only what the existential says about it, and the conclusion must not mention it. Then the argument works for the witness, whichever integer it is, because it never used anything else about it.

5.6.2 The rule of existential elimination

To use "there exists", give the witness a new name, assume only what the existential says about it, and conclude something that does not mention it.

Definition 5.6.1. Rules for the Existential Quantifier

h	$\exists x\,\varphi$						
i	x_0 | $\varphi[x_0/x]$	asm, for $\exists$E		i	$\varphi[t/x]$		
	⋮			k	$\exists x\,\varphi$	$\exists$I i	
j	χ						
k	χ	$\exists$E h, i–j					

The rule on the left is $\exists$E (**existential elimination**). It cites a line h holding $\exists x\,\varphi$ and a **flagged box** with the name x_0, a variable, whose first line is the assumption $\varphi[x_0/x]$ and whose last line j, a line of that box itself, is χ; line h and the box must be available at line

k (Definition 1.4.2, with the flagged boxes of Section 5.4). The new line k repeats χ. Side condition, at line k: the name x_0 does **not appear**

- in χ,
- in $\exists x\,\varphi$,
- in any premise, or
- in any assumption of a box **still open at line** k.

The rule on the right is ∃I, recalled from Definition 5.5.1: t is any term free for x in φ.

The box of ∃E is the flagged box of Section 5.4 with one difference: it **opens with an assumption**, $\varphi[x_0/x]$, "x_0 is a witness". The flagged box of ∀I had no assumption, because "let x_0 be arbitrary" assumes nothing about x_0. The citation "∃E h, i–j" names the line of the existential and the box, as ∨E names its disjunction and its boxes. "Appear" is the word of Section 5.2: written anywhere in the formula, after a quantifier symbol included. In a finished derivation the box's assumption is printed "asm, for ∃E"; read it as "choose x_0 with ...".

In practice, two habits make every clause of the side condition hold: name the witness with a brand-new letter, written nowhere in the derivation so far, and make sure the box's last line does not mention it. Clauses 3–5 below, the side condition, are the official test, the one a checker runs.

This rule has conditions that matter; read each one.

1. **The assumption.** The box assumes $\varphi[x_0/x]$: the formula of the existential, with the new name in place of x. This is "choose x_0 with φ". It is the **only** thing the box knows about x_0.
2. **The last line.** The box ends with χ, a line of the box itself, not of a box inside it, as for ∨E and ∀I. It is "whatever the witness is, χ follows". The new line k then asserts χ outside the box, where x_0 is no longer assumed to be anything.
3. **x_0 does not appear in χ.** The box shows that χ follows from "x_0 is a witness"; if χ mentioned x_0, line k would make a claim about a name that, outside the box, is tied to nothing. The subsection "Three ways to break the rule" below shows the damage.
4. **x_0 does not appear in a premise or in an assumption of an open box.** The name must be new to everything we know at line k. If a premise or an open assumption already said something about x_0, the assumption $\varphi[x_0/x]$ would add to it a claim that x_0 **is** the witness, which the existential never made.
5. **x_0 does not appear in $\exists x\,\varphi$.** The existential must not itself mention the name chosen for its witness. This clause also makes x_0 free for x in φ (Section 5.2: a variable that does not appear in φ is free for every variable there), so the assumption $\varphi[x_0/x]$ says of x_0 exactly what φ says of x.
6. **"Still open at line k".** The box's own assumption $\varphi[x_0/x]$ is closed at line k, so it is not among the assumptions checked. This matters: read inside the box, "an open assumption" would include $\varphi[x_0/x]$, which contains x_0 whenever x is free in φ, and the rule could then never be applied in the cases that matter.

5.6.2.1 A first witness: $n+2$ is even

Here is the proof above as a derivation, over $\mathbb{Z}$, with "n is even" written $\exists k\,(n = 2 \cdot k)$ and the computation as one line justified "arith" (Section 5.3). The premise is an existential, so its box opens first, aimed at the goal (left); the goal is postponed until a line shows its witness (right).

1	$\exists k\,(n = 2 \cdot k)$		
2	k_0 | $n = 2 \cdot k_0$		asm, for ∃E
	| ⋮		
4	| $\exists k\,(n + 2 = 2 \cdot k)$		?
5	$\exists k\,(n + 2 = 2 \cdot k)$		?

1	$\exists k\,(n = 2 \cdot k)$		
2	k_0 | $n = 2 \cdot k_0$		asm, for ∃E
3	| $n + 2 = 2 \cdot (k_0 + 1)$		arith 2
4	| $\exists k\,(n + 2 = 2 \cdot k)$		?
5	$\exists k\,(n + 2 = 2 \cdot k)$		?

1	$\exists k\,(n = 2 \cdot k)$		
2	k_0 | $n = 2 \cdot k_0$		asm, for ∃E
3	| $n + 2 = 2 \cdot (k_0 + 1)$		arith 2
4	| $\exists k\,(n + 2 = 2 \cdot k)$		∃I 3
5	$\exists k\,(n + 2 = 2 \cdot k)$		∃E 1, 2–4

At line 5, k_0 is not in line 4, not in the existential of line 1, which is also the only premise, and no box is open. The derivation and the proof match sentence for sentence:

Line	Sentence
1	"Since n is even, . . . "
box k_0, 2	". . . choose an integer k with $n = 2k$."
3	"Then $n + 2 = 2k + 2 = 2(k + 1)$, . . . " (arith)
4	". . . so $n + 2$ is even." (∃I, the witness $k + 1$)
5	(unwritten: the conclusion does not mention k, so the box closes)

Quick check. Could line 5 have been $n + 2 = 2 \cdot (k_0 + 1)$, closing the box one line earlier?

Solution. No: that line mentions the witness k_0, which breaks clause 3. Outside the box k_0 names nothing; ∃I must forget it first, as line 4 does.

5.6.2.2 One box instead of many cases

The long "or" of Section 5.1 shows why the rule has this shape. Over a domain whose elements are named m and n, $\exists x\,P(x)$ says $P(m) \vee P(n)$, and ∨E (Definition 1.6.1) would use it by proving the goal twice, once in each case. The cases look alike, so they can be done once, with a name that could be either.

1	$P(m) \lor P(n)$	
2	$\forall x\,(P(x) \to Q(f(x)))$	
3	$\quad P(m)$	asm, for $\lor$E
4	$\quad P(m) \to Q(f(m))$	$\forall$E 2
5	$\quad Q(f(m))$	$\to$E 4, 3
6	$\quad \exists y\, Q(y)$	$\exists$I 5
7	$\quad P(n)$	asm, for $\lor$E
8	$\quad P(n) \to Q(f(n))$	$\forall$E 2
9	$\quad Q(f(n))$	$\to$E 8, 7
10	$\quad \exists y\, Q(y)$	$\exists$I 9
11	$\exists y\, Q(y)$	$\lor$E 1, 3–6, 7–10

1	$\exists x\, P(x)$		
2	$\forall x\,(P(x) \to Q(f(x)))$		
3	x_0	$P(x_0)$	asm, for $\exists$E
4		$P(x_0) \to Q(f(x_0))$	$\forall$E 2
5		$Q(f(x_0))$	$\to$E 4, 3
6		$\exists y\, Q(y)$	$\exists$I 5
7	$\exists y\, Q(y)$		$\exists$E 1, 3–6

On the left, Section 1.6's rule over the two named elements; on the right, the same argument by $\exists$E. Each case box on the left is the box on the right with m or n in place of x_0. Because x_0 is fresh, one box serves for every element, named or not, and for domains that cannot be listed.

5.6.3 Using an existential: open the box early

When a resource is $\exists x\,\varphi$, the premise-directed step of the search method is to open its $\exists$E box at once, with a fresh name x_0, and to make the box's last line the **current goal**. Inside the box, the assumption $\varphi[x_0/x]$ is a new resource, and the name x_0 is a new term for later steps to use.

Name the witness early: open the $\exists$E *box as soon as the existential is available, and aim it at the current goal.*

Why early? Opening early keeps every resource, since every line available before the box is available inside it; and a box aimed at the current goal can always be closed once the goal is reached, since its name was chosen fresh. What the early box gives is the witness's name, often exactly the term a postponed $\exists$ goal is waiting for.

↻ **Second reading** · *Why opening early is safe*

And a box aimed at the current goal χ can always be closed once χ is reached: x_0 is chosen,

when the box opens, to appear in no premise, in no open assumption, not in $\exists x\,\varphi$ and not in the current goal χ. The assumptions still open at the line that closes the box are the ones that were open when it opened, so every clause of the side condition holds there. (One caution: a $\forall$I inside the box needs a flag that does not appear in the new assumption; a flag that is fresh at that point has this property.)

The first result is the derivation that this step writes almost by itself.

Proposition 5.6.1. Some Distributes Over And, One Way

$\exists x\,(P(x) \wedge Q(x)) \vdash \exists x\,P(x) \wedge \exists x\,Q(x)$.

Idea. The premise is an existential, so open its box first, with a fresh x_0, and aim it at the goal. Inside, the assumption $P(x_0) \wedge Q(x_0)$ splits by $\wedge$E, and each half is an instance of one conjunct of the goal, with the witness x_0. The goal does not mention x_0, so the box closes.

1		$\exists x\,(P(x) \wedge Q(x))$	
2	x_0	$P(x_0) \wedge Q(x_0)$	asm, for $\exists$E
		⋮	
7		$\exists x\,P(x) \wedge \exists x\,Q(x)$	?
8		$\exists x\,P(x) \wedge \exists x\,Q(x)$	?

1		$\exists x\,(P(x) \wedge Q(x))$	
2	x_0	$P(x_0) \wedge Q(x_0)$	asm, for $\exists$E
3		$P(x_0)$	$\wedge$E 2
4		$\exists x\,P(x)$	?
5		$Q(x_0)$	$\wedge$E 2
6		$\exists x\,Q(x)$	?
7		$\exists x\,P(x) \wedge \exists x\,Q(x)$	$\wedge$I 4, 6
8		$\exists x\,P(x) \wedge \exists x\,Q(x)$	?

Proof. The following is a derivation.

1		$\exists x\,(P(x) \wedge Q(x))$	
2	x_0	$P(x_0) \wedge Q(x_0)$	asm, for $\exists$E
3		$P(x_0)$	$\wedge$E 2
4		$\exists x\,P(x)$	$\exists$I 3
5		$Q(x_0)$	$\wedge$E 2
6		$\exists x\,Q(x)$	$\exists$I 5
7		$\exists x\,P(x) \wedge \exists x\,Q(x)$	$\wedge$I 4, 6
8		$\exists x\,P(x) \wedge \exists x\,Q(x)$	$\exists$E 1, 2–7

Line 2 is the assumption $\varphi[x_0/x]$ for $\varphi = P(x) \wedge Q(x)$. Lines 4 and 6 are $\exists$I with the witness x_0. At line 8, x_0 does not appear in line 7, in line 1 (which is also the only premise), or in any assumption of an open box (none is open). So line 8 is a correct $\exists$E, and the derivation has 8 lines.

■

Notice where the witness went. It was used twice inside the box, as the witness of two different existentials, and then forgotten twice. The conclusion records that something has P and something has Q, and no longer that it is the same thing.

Quick check. Is $\exists x\,P(x) \wedge \exists x\,Q(x) \vdash \exists x\,(P(x) \wedge Q(x))$? Run the search, and say where it stalls.

Solution. The premise splits by $\wedge$E into $\exists x\,P(x)$ and $\exists x\,Q(x)$. Opening both boxes early gives the assumptions $P(x_0)$ and, for the second box, a **different** name, $Q(y_0)$: the second box is opened while the first is open, so x_0 appears in an open assumption and may not be reused. The postponed goal $\exists x\,(P(x) \wedge Q(x))$ needs one term t with $P(t) \wedge Q(t)$, and the resources speak of two names, with nothing to say that they are equal. The search stalls there. It must: over $\mathbb{Z}$, with P as "even" and Q as "odd", the premise is true and the conclusion, "some integer is both even and odd", is false. That is a counter-reading (Section 5.3).

5.6.4 Three ways to break the rule

Each of clauses 3, 4 and 5 can be seen failing, and each failure leads from premises that are true, under the reading given, to a false line. Here is clause 3, beside a correct use of the same premise; the fold after it breaks clauses 4 and 5.

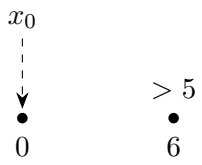
The conclusion clause (3). Over $\mathbb{N}$, "some number exceeds 5" is true. Let the $\exists$E box end in the line $x_0 > 5$, which mentions its own witness:

1	$\exists x\,(x > 5)$	
2	x_0 | $x_0 > 5$	asm, for ∃E
3	$x_0 > 5$	∃E 1, 2–2
4	x_0 | $x_0 > 5$	R 3
5	$\forall x\,(x > 5)$	∀I 4–4

1	$\exists x\,(x > 5)$	
2	x_0 | $x_0 > 5$	asm, for ∃E
3	$x_0 + 1 > 6$	arith 2
4	$\exists x\,(x > 6)$	∃I 3
5	$\exists x\,(x > 6)$	∃E 1, 2–4

Line 3 breaks clause 3. Every other line is legal: in particular line 5 is a correct ∀I, since x_0 appears in no premise and in no open assumption (there is none), and not in line 5. Over $\mathbb{N}$ the conclusion says that every number exceeds 5, which is false at 0. Once the name leaves its box, it can be generalized over as if it were arbitrary.

On the right, a correct use of the same premise: the box ends in $\exists x\,(x > 6)$, which forgets the witness before the box closes, and the conclusion is true over $\mathbb{N}$.



A small world in which the premise is true and the line that breaks clause 3 is false. The label “> 5” is the one-place predicate $x > 5$ (5 is not a dot); the dashed arrow gives x_0 the value 0. The premise is true (at 6); line 3, $x_0 > 5$, is false; so is line 5 (at 0).

Quick check. In the clause-3 derivation, line 5 is a correct ∀I, yet it is false in the picture. Which line is to blame, and why does a correct step not save the derivation?

Solution. Line 3, which closes the ∃E box with a line that mentions its witness. ∀I is safe only when the last line of its box, here line 4, is true whatever x_0 names. Line 4 is copied from line 3, and both are false when x_0 names 0. So line 5 inherits the error from line 3; it is not a new one.

Second reading · *Clauses 4 and 5, broken*

The premise clause (4). Over $\mathbb{N}$, with P read as “even” and the variable x_0 read as 1, both premises below are true: some number is even, and 1 is not even.

1	$\exists x\,P(x)$	
2	$\neg P(x_0)$	
3	x_0 | $P(x_0)$	asm, for ∃E
4	| $\bot$	¬E 3, 2
5	$\bot$	∃E 1, 3–4

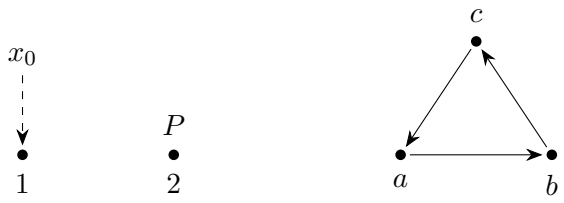
Line 5 breaks clause 4: x_0 appears in the premise on line 2. The box assumed that x_0, which

the premise had already described as odd, is the even witness, and derived a contradiction from that assumption, not from the premises.

The existential clause (5). Over $\mathbb{Z}$, with $R(x, y)$ read as "$x < y$", the premise below says that every integer has a smaller one: true.

1	$\forall y\, \exists x\, R(x, y)$	
2	$\exists x\, R(x, x_0)$	$\forall$E 1
3	x_0 | $R(x_0, x_0)$	asm, for $\exists$E
4	| $\exists z\, R(z, z)$	$\exists$I 3
5	$\exists z\, R(z, z)$	$\exists$E 2, 3–4

Line 2 is a legal $\forall$E at the term x_0. Line 5 breaks clause 5: the existential on line 2 is $\exists x\, \varphi$ with $\varphi = R(x, x_0)$, and x_0 appears in it. The assumption $\varphi[x_0/x]$ on line 3 is $R(x_0, x_0)$: the box has assumed that the witness smaller than x_0 is x_0 itself. The conclusion says that some integer is less than itself, which is false. The other clauses hold at line 5: x_0 is not in $\exists z\, R(z, z)$, not in the premise, and no box is open.



Two small worlds, one per clause: in each, the premises are true and the line that breaks the clause is false. Left, clause 4: the dots 1 *and* 2*, with* P *("even") true at* 2 *only, and* x_0 *naming* 1*; both premises,* $\exists x\, P(x)$ *and* $\neg P(x_0)$*, are true, and the line* $\bot$ *is false, as always. Right, clause 5: three dots, an arrow from* u *to* v *meaning* $R(u, v)$*; the premise* $\forall y\, \exists x\, R(x, y)$ *is true, since every dot has an arrow into it, and the line* $\exists z\, R(z, z)$ *is false, since there is no loop.*

A fourth illegal pattern, a second witness given a name that an open assumption already uses, is the error of the warning "Two witnesses, one name" below.

Warning. **The witness is not arbitrary.** The flagged boxes of $\forall$I and $\exists$E look alike: both introduce a fresh name. Their names mean opposite things. In a $\forall$I box, x_0 is **any** object, and what is proved about it holds for all. In an $\exists$E box, x_0 is **one particular** object, the witness, about which we know $\varphi[x_0/x]$ and nothing else; what is proved about it holds for it alone. From "some integer is even" nothing follows about every integer, and the rules agree:

1	$\exists x\, P(x)$	
2	x_0 | $P(x_0)$	asm, for $\exists$E
3	| x_0 | $P(x_0)$	R 2
4	| $\forall x\, P(x)$	$\forall$I 3–3
5	$\forall x\, P(x)$	$\exists$E 1, 2–4

Line 4 is blocked: x_0 appears in the assumption $P(x_0)$ of the ∃E box, which is still open at line 4. That clause of ∀I is what keeps a witness from being generalized over inside its box; clause 3 of ∃E, that x_0 not appear in χ, keeps it from leaving the box (the derivation for clause 3 above).

Here are three more derivations, each with an Idea saying why each rule was chosen: two for part (a), one for each direction, and one for part (b).

Example. Derivations with existential elimination

Show that:

(a) $\exists x\,(P(x) \lor Q(x)) \dashv\vdash \exists x\,P(x) \lor \exists x\,Q(x)$;

(b) $\exists x\,P(x),\ \forall x\,(P(x) \to \neg Q(x)) \vdash \neg\forall x\,Q(x)$.

Solution.

(a) ($\vdash$) *Idea.* The premise is an existential: open its box first. Inside, the assumption $P(x_0) \lor Q(x_0)$ is a disjunction, and the premise-directed step for ∨ (Section 1.6) opens two cases, each ending in the current goal. In each case one side of the goal has an instance with the witness x_0, so ∃I and ∨I finish it.

1		$\exists x\,(P(x) \lor Q(x))$	
2	x_0	$P(x_0) \lor Q(x_0)$	asm, for ∃E
		⋮	
9		$\exists x\,P(x) \lor \exists x\,Q(x)$	?
10		$\exists x\,P(x) \lor \exists x\,Q(x)$	?

1		$\exists x\,(P(x) \lor Q(x))$	
2	x_0	$P(x_0) \lor Q(x_0)$	asm, for ∃E
3		$P(x_0)$	asm, for ∨E
		⋮	
5		$\exists x\,P(x) \lor \exists x\,Q(x)$	?
6		$Q(x_0)$	asm, for ∨E
		⋮	
8		$\exists x\,P(x) \lor \exists x\,Q(x)$	?
9		$\exists x\,P(x) \lor \exists x\,Q(x)$	∨E 2, 3–5, 6–8
10		$\exists x\,P(x) \lor \exists x\,Q(x)$	?

1		$\exists x\,(P(x) \lor Q(x))$	
2	x_0	$P(x_0) \lor Q(x_0)$	asm, for $\exists$E
3		$P(x_0)$	asm, for $\lor$E
4		$\exists x\,P(x)$	$\exists$I 3
5		$\exists x\,P(x) \lor \exists x\,Q(x)$	$\lor$I 4
6		$Q(x_0)$	asm, for $\lor$E
7		$\exists x\,Q(x)$	$\exists$I 6
8		$\exists x\,P(x) \lor \exists x\,Q(x)$	$\lor$I 7
9		$\exists x\,P(x) \lor \exists x\,Q(x)$	$\lor$E 2, 3–5, 6–8
10		$\exists x\,P(x) \lor \exists x\,Q(x)$	$\exists$E 1, 2–9

Line 9 is $\lor$E (Definition 1.6.1), and it is the last line of the $\exists$E box itself. At line 10, x_0 is in neither line 9 nor line 1, and no box is open.

($\dashv$) *Idea.* Now the premise is a disjunction: $\lor$E first, with two cases, each ending in the goal. In each case the resource is an existential, so its box opens at once, and inside it $\lor$I and $\exists$I build the goal.

1	$\exists x\,P(x) \lor \exists x\,Q(x)$	
2	$\exists x\,P(x)$	asm, for $\lor$E
	⋮	
6	$\exists x\,(P(x) \lor Q(x))$	?
7	$\exists x\,Q(x)$	asm, for $\lor$E
	⋮	
11	$\exists x\,(P(x) \lor Q(x))$	?
12	$\exists x\,(P(x) \lor Q(x))$	$\lor$E 1, 2–6, 7–11

1	$\exists x\, P(x) \lor \exists x\, Q(x)$	
2	$\quad \exists x\, P(x)$	asm, for $\lor$E
3	$\quad x_0 \mid P(x_0)$	asm, for $\exists$E
	$\quad\quad \vdots$	
5	$\quad\quad \exists x\,(P(x) \lor Q(x))$	?
6	$\quad \exists x\,(P(x) \lor Q(x))$	$\exists$E 2, 3–5
7	$\quad \exists x\, Q(x)$	asm, for $\lor$E
8	$\quad x_0 \mid Q(x_0)$	asm, for $\exists$E
	$\quad\quad \vdots$	
10	$\quad\quad \exists x\,(P(x) \lor Q(x))$	?
11	$\quad \exists x\,(P(x) \lor Q(x))$	$\exists$E 7, 8–10
12	$\exists x\,(P(x) \lor Q(x))$	$\lor$E 1, 2–6, 7–11

1	$\exists x\, P(x) \lor \exists x\, Q(x)$	
2	$\quad \exists x\, P(x)$	asm, for $\lor$E
3	$\quad x_0 \mid P(x_0)$	asm, for $\exists$E
4	$\quad\quad P(x_0) \lor Q(x_0)$	$\lor$I 3
5	$\quad\quad \exists x\,(P(x) \lor Q(x))$	$\exists$I 4
6	$\quad \exists x\,(P(x) \lor Q(x))$	$\exists$E 2, 3–5
7	$\quad \exists x\, Q(x)$	asm, for $\lor$E
8	$\quad x_0 \mid Q(x_0)$	asm, for $\exists$E
9	$\quad\quad P(x_0) \lor Q(x_0)$	$\lor$I 8
10	$\quad\quad \exists x\,(P(x) \lor Q(x))$	$\exists$I 9
11	$\quad \exists x\,(P(x) \lor Q(x))$	$\exists$E 7, 8–10
12	$\exists x\,(P(x) \lor Q(x))$	$\lor$E 1, 2–6, 7–11

At line 6, the open assumption is line 2, which does not contain x_0; at line 11 it is line 7, which does not either. The name x_0 is used in both cases. That is allowed: when the second box opens, the first has closed, and the side condition looks only at the premises, the boxes still open, the existential and the conclusion. Lines 6 and 11 are lines of the two case boxes themselves, as $\lor$E requires.

The two derivations mirror each other: $\exists$E outside and $\lor$E inside in one, the other way round in the other. Over a domain, both sides say that some element has P or Q. Contrast Proposition 5.6.1, where only one direction holds: $\exists$ passes through $\lor$ in both directions and through $\land$ in one.

(b) *Idea.* The goal is a negation, so the goal-directed step of Section 2.2 opens a $\neg$I box on $\forall x\, Q(x)$ (Definition 2.2.1), with the goal $\bot$. Then the premise $\exists x\, P(x)$ opens its box, aimed at $\bot$. Inside, the witness x_0 is the term at which to instantiate both universal statements: the premise gives $\neg Q(x_0)$, and the assumption gives $Q(x_0)$, a clash.

1	$\exists x\, P(x)$	
2	$\forall x\,(P(x) \to \neg Q(x))$	
3	$\quad \forall x\, Q(x)$	asm, for ¬I
	$\quad \vdots$	
9	$\quad \bot$	?
10	$\neg\forall x\, Q(x)$	¬I 3–9

1	$\exists x\, P(x)$	
2	$\forall x\,(P(x) \to \neg Q(x))$	
3	$\quad \forall x\, Q(x)$	asm, for ¬I
4	$\quad x_0 \mid P(x_0)$	asm, for ∃E
	$\quad\quad \vdots$	
8	$\quad\quad \bot$	?
9	$\quad \bot$	∃E 1, 4–8
10	$\neg\forall x\, Q(x)$	¬I 3–9

1	$\exists x\, P(x)$	
2	$\forall x\,(P(x) \to \neg Q(x))$	
3	$\quad \forall x\, Q(x)$	asm, for ¬I
4	$\quad x_0 \mid P(x_0)$	asm, for ∃E
5	$\quad\quad P(x_0) \to \neg Q(x_0)$	∀E 2
6	$\quad\quad \neg Q(x_0)$	→E 5, 4
7	$\quad\quad Q(x_0)$	∀E 3
8	$\quad\quad \bot$	¬E 7, 6
9	$\quad \bot$	∃E 1, 4–8
10	$\neg\forall x\, Q(x)$	¬I 3–9

At line 9, the box of line 3 is still open, so its assumption $\forall x\, Q(x)$ is checked: x_0 does not appear in it. Nor is x_0 in $\bot$, in line 1, or in line 2. Read over $\mathbb{Z}$ with P as "odd" and Q as "divisible by 4": some integer is odd, no odd integer is divisible by 4, so not every integer is divisible by 4. The witness, whichever odd integer it is, is a counterexample to "every integer is divisible by 4", and the derivation never needs to know which.

5.6.5 Moving one quantifier past another

Section 5.1 read $\exists y \forall x\, \varphi$ and $\forall x \exists y\, \varphi$ as games, and observed that the first gives the second: if one y works for every x, then for every x some y works, namely the same one. Here is that observation as a derivation, with the letters of a two-place predicate. It shows, written out in full, a pattern that later derivations of the chapter use again and again and that later exercises

leave more and more of to the reader.

Proposition 5.6.2. Some-All Gives All-Some

$\exists x\, \forall y\, R(x,y) \vdash \forall y\, \exists x\, R(x,y)$.

Idea. Three moves, read off the goal and the premise.

① **Fix.** The goal is $\forall y$: open a flagged box with a fresh y_0 (Section 5.4); the new goal is $\exists x\, R(x, y_0)$.

② **Obtain.** That goal is an existential: postpone it. The premise is an existential: open its box early, with a fresh x_0, assuming $\forall y\, R(x_0, y)$, and aim it at the current goal $\exists x\, R(x, y_0)$. Instantiate the assumption at the term the goal mentions, y_0: this gives $R(x_0, y_0)$.

③ **Choose.** That line is an instance of the postponed goal, with the witness x_0: $\exists$I.

In the Givens/Goal table:

Givens	Goal
$\exists x\, \forall y\, R(x,y)$	$\forall y\, \exists x\, R(x,y)$
+ the fresh name y_0 (fix: $\forall$I opens a box)	$\exists x\, R(x,y_0)$: postponed
+ the witness x_0 and $\forall y\, R(x_0,y)$ (obtain: $\exists$E opens a box)	$\exists x\, R(x,y_0)$
+ $R(x_0,y_0)$ ($\forall$E at y_0)	$\exists x\, R(x,y_0)$ ✓ by $\exists$I, the witness x_0 (choose)

1	$\exists x\, \forall y\, R(x,y)$		1	$\exists x\, \forall y\, R(x,y)$		
	y_0 ⋮		2	y_0	x_0 $\forall y\, R(x_0,y)$	asm, for $\exists$E
5	$\exists x\, R(x,y_0)$	?	4		⋮	
6	$\forall y\, \exists x\, R(x,y)$	?	5		$\exists x\, R(x,y_0)$	?
			6		$\exists x\, R(x,y_0)$	$\exists$E 1, 2–4
				$\forall y\, \exists x\, R(x,y)$		?

Before reading on: at the line that closes the $\exists$E box, check each clause of its side condition. Where could x_0 still appear?

Proof. The following is a derivation.

1	$\exists x\, \forall y\, R(x,y)$			
2	y_0 | x_0	$\forall y\, R(x_0,y)$		asm, for $\exists$E
3		$R(x_0,y_0)$		$\forall$E 2
4		$\exists x\, R(x,y_0)$		$\exists$I 3
5		$\exists x\, R(x,y_0)$		$\exists$E 1, 2–4
6	$\forall y\, \exists x\, R(x,y)$			$\forall$I 2–5

Line 2 is the assumption $\varphi[x_0/x]$ for $\varphi = \forall y\, R(x,y)$. Line 3 is $\forall$E at y_0. Line 4 is $\exists$I with $\varphi = R(x,y_0)$ and the witness x_0, which is free for x since $R(x,y_0)$ has no quantifier.

Line 5 closes the $\exists$E box, whose last line, line 4, is a line of the box itself. Its side condition holds: x_0 does not appear in $\chi = \exists x\, R(x,y_0)$, nor in the existential $\exists x\, \forall y\, R(x,y)$ of line 1, which is also the only premise; and the one box still open at line 5, the flagged box of y_0, has no assumption.

Line 6 closes the flagged box of y_0 by $\forall$I (Definition 5.4.1). That box immediately contains the $\exists$E box, so its range starts at line 2 (the convention of Section 5.4), and its last line, line 5, is a line of the box itself, the instance $(\exists x\, R(x,y))[y_0/y]$. The name y_0 appears neither in the premise nor in line 6, and at line 6 no box is open. This proves the proposition.

■

The two names played opposite parts. The name y_0 is arbitrary: nothing was assumed about it, and line 6 generalizes over it. The name x_0 is the witness: line 2 assumed everything the premise says about it, and line 5 forgets it. The witness x_0 works for y_0 because it works for every y, which line 2 says. This is the shape of a great many proofs in analysis, and it has a name.

Fix, obtain, choose: fix the arbitrary objects of the goal ($\forall$I); obtain witnesses from the hypotheses ($\forall$E and $\exists$E); choose the goal's witness from what is now available ($\exists$I).

Nothing in the proof used that the formula is $R(x,y)$: the same six lines derive $\exists x\, \forall y\, \varphi \vdash \forall y\, \exists x\, \varphi$ for any formula φ, which is the derivation promised in Section 5.1; the fold checks the substitutions.

Second reading · *The swap for any formula*

For any formula φ and different variables x and y, with different names x_0 and y_0, other than x and y, that do not appear in φ, the same six lines derive $\exists x\, \forall y\, \varphi \vdash \forall y\, \exists x\, \varphi$. Line 2 is $\forall y\, (\varphi[x_0/x])$. Line 3 is $(\varphi[x_0/x])[y_0/y]$, which equals $(\varphi[y_0/y])[x_0/x]$, since the two substitutions put fresh names in for two different variables and so do not interfere; so line 4 is $\exists$I with the formula $\varphi[y_0/y]$ and the witness x_0. The fresh names are free for their variables, since neither appears in φ, and the side conditions are checked as before. With the letters x and y exchanged, this is the derivation of $\exists y\, \forall x\, \varphi$ to $\forall x\, \exists y\, \varphi$ promised in Section 5.1.

↻ **Second reading** · *Either nesting works*

Either nesting works. The derivation opened the flagged box of y_0 first and the ∃E box inside it. The other order is just as good.

1	$\exists x\, \forall y\, R(x, y)$	
2	x_0 | $\forall y\, R(x_0, y)$	asm, for ∃E
3	y_0 | $R(x_0, y_0)$	∀E 2
4	$\exists x\, R(x, y_0)$	∃I 3
5	$\forall y\, \exists x\, R(x, y)$	∀I 3–4
6	$\forall y\, \exists x\, R(x, y)$	∃E 1, 2–5

At line 5, y_0 appears neither in the premise, nor in the open assumption $\forall y\, R(x_0, y)$ (its y is a bound variable, not y_0), nor in line 5. At line 6, x_0 is not in line 5, not in line 1, and no box is open. In this order the ∃E box is aimed at the whole goal; in the first, at the goal left inside the flagged box. Both are the search method's "open the box early": early relative to the goal that the box must reach.

The converse search stalls at a side condition. Try $\forall y\, \exists x\, R(x, y) \vdash \exists x\, \forall y\, R(x, y)$. The goal is an existential: postpone it; it will need a witness x_0 and a line $\forall y\, R(x_0, y)$. The premise is a universal, and to get an existential out of it we must instantiate it at some term. The natural routes then meet a side condition. Here are two of them.

First route: obtain the witness, then try to generalize. Instantiate at a name y_0, open the ∃E box on the result, and try to prove $\forall y\, R(x_0, y)$ inside it by ∀I:

1	$\forall y\, \exists x\, R(x, y)$	
2	$\exists x\, R(x, y_0)$	∀E 1
3	x_0 | $R(x_0, y_0)$	asm, for ∃E
4	y_0 | $R(x_0, y_0)$	R 3
5	$\forall y\, R(x_0, y)$	∀I 4–4
6	$\exists x\, \forall y\, R(x, y)$	∃I 5
7	$\exists x\, \forall y\, R(x, y)$	∃E 2, 3–6

Line 5 is blocked: y_0 appears in the assumption $R(x_0, y_0)$ of the ∃E box, which is still open. The witness x_0 was obtained **for** y_0, and the assumption says so. Choosing a different flag, say z_0, does not help: the goal inside its box would be $R(x_0, z_0)$, and nothing gives it.

Second route: fix first, obtain inside, and carry the witness out. Open the flagged box of y_0 first, obtain x_0 inside it, and try to close the ∃E box with a line about x_0, so that ∀I can then generalize over y_0:

1	$\forall y\,\exists x\,R(x,y)$	
2	$y_0 \mid \exists x\,R(x,y_0)$	$\forall$E 1
3	$\quad x_0 \mid R(x_0,y_0)$	asm, for $\exists$E
4	$\quad R(x_0,y_0)$	$\exists$E 2, 3–3
5	$\forall y\,R(x_0,y)$	$\forall$I 2–4
6	$\exists x\,\forall y\,R(x,y)$	$\exists$I 5

Line 4 is blocked: x_0 appears in $\chi = R(x_0,y_0)$. Lines 5 and 6 would be legal on their own; the derivation fails only because the witness escaped its box at line 4.

The two blocked lines are the two halves of one fact. The witness x_0 is obtained after y_0 is fixed, so it may depend on y_0; the conclusion asks for an x_0 chosen before any y, which may not. This is the dependence that Section 5.1 read off the order of the quantifiers, now visible as the order of the boxes. A stalled search is not a proof that no derivation exists. A counter-reading breaks the claim: two dots 0 and 1, with $R(x,y)$ read as "x differs from y". Each dot has a dot that differs from it, so the premise is true; no dot differs from both dots, itself included, so the conclusion is false.

Second reading · *Renaming the bound variable of an existential*

Renaming the bound variable of an existential. Section 5.2 promised that a bound variable at the front of a formula may be renamed, and Section 5.4 derived the renaming for $\forall$ (the fold after Example (Flagged boxes at work)). Here is the derivation for $\exists$. Let φ be a formula and y, z variables, with z not appearing in φ. Then

$$\exists y\,\varphi \;\dashv\vdash\; \exists z\,\varphi[z/y].$$

Take a fresh name z_0, a variable different from y and z that does not appear in φ, as in Section 5.4. Both directions use one fact about substitution: since z does not appear in φ, each free z of $\varphi[z/y]$ stands where φ had a free y, and nowhere else, so putting a term t in for those z's gives $\varphi[t/y]$; that is, $(\varphi[z/y])[t/z] = \varphi[t/y]$. (It is a structural induction of the kind in Section 5.2, one clause at a time.)

1	$\exists y\,\varphi$	
2	$z_0 \mid \varphi[z_0/y]$	asm, for $\exists$E
3	$\quad \exists z\,\varphi[z/y]$	$\exists$I 2
4	$\exists z\,\varphi[z/y]$	$\exists$E 1, 2–3

1	$\exists z\,\varphi[z/y]$	
2	$z_0 \mid \varphi[z_0/y]$	asm, for $\exists$E
3	$\quad \exists y\,\varphi$	$\exists$I 2
4	$\exists y\,\varphi$	$\exists$E 1, 2–3

In the left derivation, line 3 is $\exists$I with the formula $\varphi[z/y]$, the variable z and the witness z_0: by the fact, $(\varphi[z/y])[z_0/z]$ is line 2. The witness is free for z, because the quantifiers of $\varphi[z/y]$ are those of φ, and none of them is on z_0. In the right derivation, line 2 is the assumption of the box for the existential $\exists z\,\varphi[z/y]$, which is $(\varphi[z/y])[z_0/z]$, that is $\varphi[z_0/y]$ by the fact, and line 3 is $\exists$I with φ, y and the witness z_0, which is free for y since it does not appear in φ. In both, line 4 is a correct $\exists$E: z_0 does not appear in φ, so it is in neither line 1 nor line 3, and no box is open.

So $\exists y\,(y \neq x)$ and $\exists z\,(z \neq x)$ can each be derived from the other, as the recipe of Section 5.2

assumed.

5.6.6 Two witnesses

A proof often obtains several witnesses, one from each existential hypothesis. Each needs its own box and its own name.

Example. The sum of two even integers is even

Over $\mathbb{Z}$, with "n is even" written $\exists k\,(n = 2 \cdot k)$, derive

$$\exists k\,(n = 2 \cdot k),\ \exists k\,(m = 2 \cdot k) \ \vdash\ \exists k\,(n + m = 2 \cdot k).$$

Solution. Idea. Both premises are existentials: open both boxes early, the second inside the first, with **different** names k_0 and l_0. The goal is an existential: postpone it. Inside both boxes the resources are $n = 2 \cdot k_0$ and $m = 2 \cdot l_0$, and arithmetic gives $n + m = 2 \cdot (k_0 + l_0)$, which shows the witness: the compound term $k_0 + l_0$.

1		$\exists k\,(n = 2 \cdot k)$	
2		$\exists k\,(m = 2 \cdot k)$	
3	k_0	$n = 2 \cdot k_0$	asm, for ∃E
		⋮	
7		$\exists k\,(n + m = 2 \cdot k)$	?
8		$\exists k\,(n + m = 2 \cdot k)$	?

1			$\exists k\,(n = 2 \cdot k)$	
2			$\exists k\,(m = 2 \cdot k)$	
3	k_0		$n = 2 \cdot k_0$	asm, for ∃E
4		l_0	$m = 2 \cdot l_0$	asm, for ∃E
			⋮	
6			$\exists k\,(n + m = 2 \cdot k)$	?
7			$\exists k\,(n + m = 2 \cdot k)$	∃E 2, 4–6
8			$\exists k\,(n + m = 2 \cdot k)$	?

1		$\exists k\,(n = 2 \cdot k)$	
2		$\exists k\,(m = 2 \cdot k)$	
3	k_0	$n = 2 \cdot k_0$	asm, for ∃E
4	k_0 l_0	$m = 2 \cdot l_0$	asm, for ∃E
5	k_0 l_0	$n + m = 2 \cdot (k_0 + l_0)$	arith 3, 4
6	k_0 l_0	$\exists k\,(n + m = 2 \cdot k)$	∃I 5
7	k_0	$\exists k\,(n + m = 2 \cdot k)$	∃E 2, 4–6
8		$\exists k\,(n + m = 2 \cdot k)$	∃E 1, 3–7

Line 5 is one line of arithmetic: from $n = 2k_0$ and $m = 2l_0$, substituting and factoring gives $n + m = 2k_0 + 2l_0 = 2(k_0 + l_0)$. The substitutions are steps of the rule =E, which arrives only in Section 5.7, so here the whole calculation is one line justified "arith", citing the two lines it uses. Line 6 is ∃I with $\varphi = (n + m = 2 \cdot k)$ and the witness $k_0 + l_0$.

At line 7 the only box still open is the outer one, whose assumption is line 3; l_0 appears neither in it, nor in the premises on lines 1 and 2 (line 2 is also the existential cited), nor in line 6. At line 8, k_0 appears neither in line 7, nor in the premises (line 1 is the existential cited), and no box is open. The derivation has 8 lines. The written proof is three sentences: *Choose k with $n = 2k$, and l with $m = 2l$. Then $n + m = 2(k + l)$. So $n + m$ is even.*

Quick check. Why must the second witness in Example (The sum of two even integers is even) be called l_0, and not k_0 again? Name the line and the clause of the side condition that a second k_0 would break.

Solution. The inner box would then be flagged k_0 with the assumption $m = 2 \cdot k_0$, and line 7 would close it by ∃E. At line 7 the outer box is still open, and its assumption $n = 2 \cdot k_0$ (line 3) contains k_0. That breaks the clause "x_0 does not appear in any assumption of a box still open at line k". In words: the second witness would be assumed to be the same integer as the first, so n and m would be forced equal, which the premises never said.

5.6.7 A contradiction trapped in a box

An ∃E box may end in ⊥. The contradiction is found inside the box, about the witness, and leaves it as ⊥, which mentions no name. The best-known instance is a theorem of pure logic.

Theorem 5.6.3. The Barber

For every two-place predicate symbol R,

$$\vdash \neg\exists y\,\forall x\,(R(x, y) \leftrightarrow \neg R(x, x)).$$

Try it first (five minutes): the goal is a negation.

Idea. The goal is a negation: open a ¬I box on the existential, aimed at $\bot$. Its assumption is an existential: name the witness early, y_0, aiming the ∃E box at the same $\bot$. Inside, an attempt that instantiates the universal at some other term t gets $R(t, y_0) \leftrightarrow \neg R(t, t)$, and nothing clashes with it; the one term that makes both sides speak of the same thing is y_0 itself. The instance says that $R(y_0, y_0)$ holds exactly when it does not, and a statement equivalent to its own negation is contradictory (Lemma 2.4.2).

1	$\exists y \forall x \, (R(x, y) \leftrightarrow \neg R(x, x))$	asm, for ¬I
	⋮	
10	$\bot$	?
11	$\neg \exists y \forall x \, (R(x, y) \leftrightarrow \neg R(x, x))$	?

1	$\exists y \forall x \, (R(x, y) \leftrightarrow \neg R(x, x))$		asm, for ¬I
2	y_0	$\forall x \, (R(x, y_0) \leftrightarrow \neg R(x, x))$	asm, for ∃E
		⋮	
9		$\bot$	?
10	$\bot$		∃E 1, 2–9
11	$\neg \exists y \forall x \, (R(x, y) \leftrightarrow \neg R(x, x))$		?

Proof. The following is a derivation with no premises.

1	$\exists y \forall x \, (R(x, y) \leftrightarrow \neg R(x, x))$		asm, for ¬I
2	y_0	$\forall x \, (R(x, y_0) \leftrightarrow \neg R(x, x))$	asm, for ∃E
3		$R(y_0, y_0) \leftrightarrow \neg R(y_0, y_0)$	∀E 2
4		$R(y_0, y_0)$	asm, for ¬I
5		$\neg R(y_0, y_0)$	↔E 3, 4
6		$\bot$	¬E 4, 5
7		$\neg R(y_0, y_0)$	¬I 4–6
8		$R(y_0, y_0)$	↔E 3, 7
9		$\bot$	¬E 8, 7
10	$\bot$		∃E 1, 2–9
11	$\neg \exists y \forall x \, (R(x, y) \leftrightarrow \neg R(x, x))$		¬I 1–10

Lines 4–9 are lines 2–7 of the derivation of Lemma 2.4.2 at the statement $R(y_0, y_0)$, with line 3 in place of its assumption. At line 10, y_0 is not in $\bot$, not in the existential of line 1 (its y is bound), and there is no premise; the only open assumption, line 1, does not contain y_0. This proves the theorem.

■

Read $R(x, y)$ as "y shaves x": nobody shaves exactly those who do not shave themselves. The old puzzle of the barber who does is answered by this theorem: there is no such barber. Read the domain as consisting of sets and $R(x, y)$ as $x \in y$: no set has as its elements exactly the sets that are not elements of themselves. For sets this is Russell's theorem, which *Sets, Relations, and Cardinality*, Section 1.1 proves as a theorem about sets. Neither reading was used in the proof, which holds for every R; the "paradox" is eleven lines of logic, and its conclusion is that something does not exist.

Quick check. In the proof of Theorem 5.6.3, lines 3–9 are all about y_0. Why may line 10 close the ∃E box?

Solution. Its last line, $\bot$, does not mention y_0 (clause 3); y_0 appears in no premise and in no open assumption (line 1's y is bound), and not in the existential of line 1 (clauses 4 and 5).

5.6.8 How this rule looks in a mathematical proof

In a written proof, the ∃E box is announced by "choose", "let x_0 be such a point", "say $n = 2k$", "fix such a k". The assumption of the box is the property in the sentence, and the rest of the paragraph is the inside of the box. The claim the paragraph ends on, usually the clause after the final "so", must not mention the chosen name, though the computation leading up to it may. If that claim still mentions the name, the paragraph has not finished: a later step must reach a claim without the name before the box can close, often an ∃I that forgets it (as in the example below), or a contradiction $\bot$ (as in Example (Derivations with existential elimination) (b)).

↻ **Second reading** · *A bounded sequence, rule by rule*

Here is an instance from analysis. A sequence (a_n) is **bounded** if there is an M with $|a_n| \le M$ for every $n \in \mathbb{N}$. A proof says:

> *Since (a_n) is bounded, choose M with $|a_n| \le M$ for all $n \in \mathbb{N}$. Then $|2a_n| = 2|a_n| \le 2M$ for all $n \in \mathbb{N}$, so $(2a_n)$ is bounded.*

With the domain $\mathbb{R}$, the sequence as a function symbol a, and the bound on n written out by Definition 5.1.1, the proof is this derivation.

1		$\exists M\,\forall n\,(n \in \mathbb{N} \to \lvert a_n\rvert \le M)$	
2	M_0	$\forall n\,(n \in \mathbb{N} \to \lvert a_n\rvert \le M_0)$	asm, for $\exists$E
3	M_0 n_0	$n_0 \in \mathbb{N}$	asm, for $\to$I
4	M_0 n_0	$n_0 \in \mathbb{N} \to \lvert a_{n_0}\rvert \le M_0$	$\forall$E 2
5	M_0 n_0	$\lvert a_{n_0}\rvert \le M_0$	$\to$E 4, 3
6	M_0 n_0	$\lvert 2 \cdot a_{n_0}\rvert \le 2 \cdot M_0$	arith 5
7	M_0	$n_0 \in \mathbb{N} \to \lvert 2 \cdot a_{n_0}\rvert \le 2 \cdot M_0$	$\to$I 3–6
8		$\forall n\,(n \in \mathbb{N} \to \lvert 2 \cdot a_n\rvert \le 2 \cdot M_0)$	$\forall$I 3–7
9		$\exists M\,\forall n\,(n \in \mathbb{N} \to \lvert 2 \cdot a_n\rvert \le M)$	$\exists$I 8
10		$\exists M\,\forall n\,(n \in \mathbb{N} \to \lvert 2 \cdot a_n\rvert \le M)$	$\exists$E 1, 2–9

The sentences match the boxes. "Choose M" opens the $\exists$E box (line 2). "For all $n \in \mathbb{N}$" in the second sentence is the flagged box of n_0 with its $\to$I box (lines 3–7, closed by line 8); the arithmetic $|2a_n| = 2|a_n| \le 2M$ is line 6. "So $(2a_n)$ is bounded" is two steps: $\exists$I with the witness $2 \cdot M_0$ (line 9), which forgets the bound $2M$, and the close of the box (line 10). The conclusion mentions neither M_0 nor $2 \cdot M_0$, as the side condition requires.

Warning. **Two witnesses, one name.** "Let n and m be even, say $n = 2k$ and $m = 2k$" is a common error in parity proofs. It gives two unknown witnesses the same name, and the rest of the proof then uses an equation the premises never gave. For instance it "proves" that the sum of two even integers is divisible by 4:

1		$\exists k\,(n = 2 \cdot k)$	
2		$\exists k\,(m = 2 \cdot k)$	
3	k_0	$n = 2 \cdot k_0$	asm, for $\exists$E
4	k_0 k_0	$m = 2 \cdot k_0$	asm, for $\exists$E
5	k_0 k_0	$n + m = 4 \cdot k_0$	arith 3, 4
6	k_0 k_0	$\exists k\,(n + m = 4 \cdot k)$	$\exists$I 5
7	k_0	$\exists k\,(n + m = 4 \cdot k)$	$\exists$E 2, 4–6
8		$\exists k\,(n + m = 4 \cdot k)$	$\exists$E 1, 3–7

Line 7 is blocked: k_0 appears in the assumption $n = 2 \cdot k_0$ of the outer box, still open at line 7. With $n = 2$ and $m = 4$, both premises are true and $n + m = 6$ is not divisible by 4.

5.6.9 The four quantifier rules on one card

Example. Instances of $\exists$E

The existential	The box's assumption (name)	The box's last line	The new line
$\exists k\,(m = 2 \cdot k + 1)$	$m = 2 \cdot k_0 + 1$ (k_0)	$\exists j\,(m + 1 = 2 \cdot j)$	$\exists j\,(m + 1 = 2 \cdot j)$
$\exists y\,(R(f(c), y) \wedge \forall z\,R(y, z))$	$R(f(c), u_0) \wedge \forall z\,R(u_0, z)$ (u_0)	$\exists y\,R(y, y)$	$\exists y\,R(y, y)$
$\exists \delta\,(\delta > 0 \wedge \delta < \varepsilon)$	$\delta_0 > 0 \wedge \delta_0 < \varepsilon$ (δ_0)	$\varepsilon > 0$	$\varepsilon > 0$

Inside each box the last line is reached from the assumption (in the first by arith, $m + 1 = 2 \cdot (k_0 + 1)$, and $\exists$I; in the second by $\wedge$E, $\forall$E at u_0 and $\exists$I; in the third by arith from $0 < \delta_0 < \varepsilon$), and none mentions its name.

Not an instance. From $\exists x\,R(x, c)$, a box that opens with $R(c, x_0)$: the assumption must be $\varphi[x_0/x] = R(x_0, c)$; $R(c, x_0)$ assumes the witness stands on the other side of c, which the existential never said.

Chapters 1 and 2's thirteen rules are on the card at the end of Section 3.1. Here are the four quantifier rules, with their side conditions. Section 5.7 collects the whole table, and nothing later depends on the last column.

Rule	Shape	Side condition	You write	Lean
$\forall$E	$\forall x\,\varphi \Longrightarrow \varphi[t/x]$	t free for x in φ	"In particular, for $x = t$, …"; "applying the hypothesis to t, …"	`h t` (or `specialize h t`)
$\forall$I	box $[x_0]$ … $\varphi[x_0/x] \Longrightarrow \forall x\,\varphi$	x_0 in no premise, in no open assumption, not in $\forall x\,\varphi$	"Let x be arbitrary. … Since x was arbitrary, …"	`intro x`
$\exists$I	$\varphi[t/x] \Longrightarrow \exists x\,\varphi$	t free for x in φ	"Take $x = t$. Then …"; "$2k + 1$ will do."	`use t` (or `Exists.intro t h`)
$\exists$E	$\exists x\,\varphi$, box $[x_0]$ $\varphi[x_0/x]$ … χ $\Longrightarrow \chi$	x_0 not in χ, not in $\exists x\,\varphi$, in no premise, in no open assumption	"Choose k with $n = 2k$. Then …"; "let x_0 be such a point."	`obtain` (or `Exists.elim h`)

↻ **Second reading** · *Choosing a name*

The phrase "choose" is apt and slightly misleading. We do not choose the witness, since we cannot see it; we choose a **name** for it. A written proof that says "choose k" promises the reader that nothing more about k will be used than the property stated, and that k will be gone by the end of the argument.

Warning (Not allowed: ∃E).

- "Close the box on a line that mentions its witness": clause 3 (the pair in "Three ways to break the rule").
- "Name the witness with a letter that a premise or an open assumption mentions": clause 4 (the fold "Clauses 4 and 5, broken"; the warning "Two witnesses, one name").
- "Name the witness with a letter that the existential mentions": clause 5 (the fold "Clauses 4 and 5, broken").
- "Open the box with an assumption other than $\varphi[x_0/x]$": clause 1 ("Not an instance" in Example (Instances of ∃E)).

The decision chart grows by two rows:

Main symbol	As the goal (goal-directed)	As an available line (premise-directed)
$\forall x\,\varphi$	open a flagged box with a brand-new name x_0; its last line is $\varphi[x_0/x]$; close with ∀I	instantiate at the term the goal, or another line, names (∀E)
$\exists x\,\varphi$	postpone; use ∃I once a line shows an instance $\varphi[t/x]$	open the ∃E box at once, with a brand-new witness name, aimed at the current goal

The connective rows are the table of moves in Section 3.1.

The search method: one more step. With this rule, each of the four quantifier rules has added a step.

- **Premise $\exists x\,\varphi$.** Open the ∃E box early, as soon as the existential is available, with a fresh name x_0 and the assumption $\varphi[x_0/x]$; make its last line the current goal, which must not mention x_0. Give every new witness a new name.

In the Givens/Goal table, ∃E puts a witness among the Givens, its name and its assumption $\varphi[x_0/x]$, and leaves the goal as it was.

Remark (Looking ahead). A derivation with n existential premises opens n boxes and names n witnesses, one choice per box. *Sets, Relations, and Cardinality*, Chapter 3 asks when infinitely many witnesses may be chosen at once, and answers with the axiom of choice.

5.6.10 Exercises

5.6.10.1 A. Check your understanding

Exercise. A1

These parts are about the proof of Proposition 5.6.2.

(a) State ∃E as a schema, with every clause of its side condition.

(b) The proof opened the flagged box of y_0 first and the ∃E box of x_0 inside it. Could they

be opened in the other order? If so, does every side condition still hold?

(c) Answer the *Before reading on* prompt of the proof: at line 5, where could x_0 still appear, and does it?

(d) In the two attempts at the converse, $\forall y\,\exists x\,R(x,y) \vdash \exists x\,\forall y\,R(x,y)$, which line is blocked in each, and by which clause of which rule?

(e) Give the idea of the proof in two sentences.

Solution. *(a)* From $\exists x\,\varphi$ on line h and a flagged box with name x_0, from the assumption $\varphi[x_0/x]$ on line i to χ on line j, a line of the box itself, write χ on line k, justified "$\exists$E h, i–j". Side condition at line k: x_0 does not appear in χ, in $\exists x\,\varphi$, in any premise, or in any assumption of a box still open at line k.

(b) Yes, either order works. In the other order (the fold "Either nesting works" after the proof), the $\exists$E box encloses the flagged box. At the $\forall$I line, y_0 appears in no premise, not in the open assumption $\forall y\,R(x_0,y)$, and not in the conclusion; at the $\exists$E line, x_0 is not in $\forall y\,\exists x\,R(x,y)$, not in the premise, and no box is open.

(c) At line 5, the clauses list four places: χ, which is line 4, $\exists x\,R(x,y_0)$; the existential of line 1, $\exists x\,\forall y\,R(x,y)$; the premise, which is line 1 again; and the assumptions of boxes open at line 5. The only box open at line 5 is the flagged box of y_0, which has no assumption. The name x_0 appears in none of these, so line 5 is correct. It does appear in lines 2–4, inside the box, which the side condition allows.

(d) In the first attempt, line 5 ($\forall$I over y_0) is blocked by the clause "y_0 appears in no assumption of a box still open", since the $\exists$E box's assumption $R(x_0,y_0)$ contains y_0. In the second, line 4 ($\exists$E) is blocked by the clause "x_0 does not appear in χ", since $\chi = R(x_0,y_0)$.

(e) Fix an arbitrary y_0 and obtain the witness x_0 of the premise, for which $R(x_0,y)$ holds for every y, in particular for y_0. So x_0 is a witness for $\exists x\,R(x,y_0)$, and since y_0 was arbitrary, $\forall y\,\exists x\,R(x,y)$ follows.

Exercise. A2: Drill

Each part takes a line or two. The parts run from a single step, through a missing line, a justification column and steps to judge, to short derivations. "Clause" means a clause of the reading of Definition 5.6.1.

(a) A line h is $\exists x\,P(x)$. Write the first line of its $\exists$E box, with the name u_0.

(b) Line 9 is $\exists y\,Q(y)$, justified "$\exists$E 3, 4–8", and line 4 is the assumption $P(u_0) \wedge R(u_0,c)$. What is line 3?

(c) Justify each line: 1 $\exists x\,(P(x) \wedge Q(x))$, premise; 2 $\forall x\,(Q(x) \to R(x))$, premise; in a box flagged x_0: 3 $P(x_0) \wedge Q(x_0)$; 4 $Q(x_0)$; 5 $Q(x_0) \to R(x_0)$; 6 $R(x_0)$; 7 $P(x_0)$; 8 $P(x_0) \wedge R(x_0)$; 9 $\exists x\,(P(x) \wedge R(x))$; then, outside the box, 10 $\exists x\,(P(x) \wedge R(x))$.

(d) An $\exists$E box for $\exists k\,(m = 2 \cdot k)$, named k_0, ends in $m^2 = 4 \cdot k_0^2$. May it close with that line?

(e) The premise is $Q(x_0)$. May the box of $\exists x\,P(x)$ be named x_0?

(f) Premises $\exists x\,P(x)$, $\exists x\,Q(x)$. May both boxes be named x_0 if one is inside the other?

(g) May two boxes share a name if the first has closed before the second opens?

(h) Inside an ∃E box with assumption $P(x_0)$, may ∀I generalize over x_0?

(i) Find the illegal line: 1 $\exists x\, P(x)$, premise; in a box flagged x_0: 2 $P(x_0)$, assumption; then 3 $P(x_0)$, ∃E 1, 2–2; 4 $\exists y\, P(y)$, ∃I 3.

(j) Derive $\exists x\,(P(x) \wedge Q(x)) \vdash \exists x\, Q(x)$.

(k) Derive $\exists x\, P(x),\ \forall x\,(P(x) \to Q(x)) \vdash \exists x\, Q(x)$.

(l) Derive $\exists x\, \exists y\, R(x,y) \vdash \exists y\, \exists x\, R(x,y)$.

Solution.

(a) The box flagged u_0, opening with the assumption $P(u_0)$.

(b) $\exists x\,(P(x) \wedge R(x,c))$, with any bound letter in place of x (one that does not clash with c or u_0).

(c) 3 assumption, for ∃E; 4 ∧E 3; 5 ∀E 2 (at x_0); 6 →E 5, 4; 7 ∧E 3; 8 ∧I 7, 6; 9 ∃I 8; 10 ∃E 1, 3–9.

(d) No, clause 3: the last line mentions k_0.

(e) No, clause 4: x_0 appears in a premise.

(f) No, clause 4: the inner box's name appears in the open assumption $P(x_0)$ of the outer one.

(g) Yes, as in Example (Derivations with existential elimination) (a), (⊣).

(h) No: x_0 is in an open assumption, which the second clause of ∀I forbids; the witness is not arbitrary.

(i) Line 3, clause 3: its box's last line mentions x_0. (Line 4 would be legal; the conclusion is even derivable, with the ∃I inside the box, but this derivation does not derive it.)

(j) In a box flagged x_0: 2 $P(x_0) \wedge Q(x_0)$, assumption; 3 $Q(x_0)$, ∧E 2; 4 $\exists x\, Q(x)$, ∃I 3; then 5 $\exists x\, Q(x)$, ∃E 1, 2–4.

(k) In a box flagged x_0: 3 $P(x_0)$, assumption; 4 $P(x_0) \to Q(x_0)$, ∀E 2; 5 $Q(x_0)$, →E 4, 3; 6 $\exists x\, Q(x)$, ∃I 5; then 7 $\exists x\, Q(x)$, ∃E 1, 3–6.

(l) In a box flagged x_0: 2 $\exists y\, R(x_0,y)$, assumption; inside it, in a box flagged y_0: 3 $R(x_0,y_0)$, assumption; 4 $\exists x\, R(x,y_0)$, ∃I 3; 5 $\exists y\, \exists x\, R(x,y)$, ∃I 4; then 6 $\exists y\, \exists x\, R(x,y)$, ∃E 2, 3–5; and 7, the same line, ∃E 1, 2–6.

5.6.10.2 B. Practice

Exercise. B1

Derive $\forall x\, \neg(P(x) \wedge Q(x)),\ \exists x\, P(x) \vdash \exists x\, \neg Q(x)$. Read the sequent over $\mathbb{Z}$ with P as "positive" and Q as "negative".

Solution. Idea. The premise $\exists x\, P(x)$ opens its box first, with a witness x_0, aimed at the goal. The goal $\exists x\, \neg Q(x)$ is postponed; the witness x_0 is the natural candidate, so the new goal inside is $\neg Q(x_0)$, a negation: a ¬I box on $Q(x_0)$. Inside, $P(x_0) \wedge Q(x_0)$ clashes with the first premise at x_0.

1	$\forall x\, \neg(P(x) \wedge Q(x))$	
2	$\exists x\, P(x)$	
3	x_0 | $P(x_0)$	asm, for $\exists$E
4	$Q(x_0)$	asm, for $\neg$I
5	$P(x_0) \wedge Q(x_0)$	$\wedge$I 3, 4
6	$\neg(P(x_0) \wedge Q(x_0))$	$\forall$E 1
7	$\bot$	$\neg$E 5, 6
8	$\neg Q(x_0)$	$\neg$I 4–7
9	$\exists x\, \neg Q(x)$	$\exists$I 8
10	$\exists x\, \neg Q(x)$	$\exists$E 2, 3–9

At line 10, x_0 is not in line 9, not in line 2, not in either premise, and no box is open. Over $\mathbb{Z}$: no integer is both positive and negative, and some integer is positive, so some integer is not negative. The witness of the conclusion is the positive witness of the premise.

Exercise. B2: Fix, obtain, choose, faded

Here is a derivation of

$$\forall x\, \exists y\, R(x,y),\ \forall x\, \forall y\, (R(x,y) \to Q(y)) \ \vdash\ \forall x\, \exists y\, (R(x,y) \wedge Q(y))$$

with its last line missing.

1	$\forall x\, \exists y\, R(x,y)$	
2	$\forall x\, \forall y\, (R(x,y) \to Q(y))$	
3	x_0 | $\exists y\, R(x_0,y)$	$\forall$E 1
4	y_0 | $R(x_0,y_0)$	asm, for $\exists$E
5	$\forall y\, (R(x_0,y) \to Q(y))$	$\forall$E 2
6	$R(x_0,y_0) \to Q(y_0)$	$\forall$E 5
7	$Q(y_0)$	$\to$E 6, 4
8	$R(x_0,y_0) \wedge Q(y_0)$	$\wedge$I 4, 7
9	$\exists y\, (R(x_0,y) \wedge Q(y))$	$\exists$I 8
10	$\exists y\, (R(x_0,y) \wedge Q(y))$	$\exists$E 3, 4–9

Name the move, then complete it. Then mark which lines are the "fix", which the "obtain", and which the "choose".

Solution. The missing line closes the flagged box of x_0 by ∀I; the move is the "fix" of fix, obtain, choose, completed: "since x_0 was arbitrary".

1		$\forall x\, \exists y\, R(x,y)$	
2		$\forall x\, \forall y\, (R(x,y) \to Q(y))$	
3	x_0	$\exists y\, R(x_0,y)$	∀E 1
4		y_0 | $R(x_0,y_0)$	asm, for ∃E
5			$\forall y\, (R(x_0,y) \to Q(y))$ ∀E 2
6			$R(x_0,y_0) \to Q(y_0)$ ∀E 5
7			$Q(y_0)$ →E 6, 4
8			$R(x_0,y_0) \wedge Q(y_0)$ ∧I 4, 7
9			$\exists y\, (R(x_0,y) \wedge Q(y))$ ∃I 8
10		$\exists y\, (R(x_0,y) \wedge Q(y))$	∃E 3, 4–9
11		$\forall x\, \exists y\, (R(x,y) \wedge Q(y))$	∀I 3–10

Line 11 cites the flagged box 3–10; its last line, line 10, is a line of that box itself and is the instance $\psi[x_0/x]$ for $\psi = \exists y\, (R(x,y) \wedge Q(y))$. The name x_0 appears in neither premise and not in line 11, and at line 11 no box is open. (At line 10, y_0 is not in line 9, not in line 3, not in the premises, and the only open box, that of x_0, has no assumption.)

The moves: **fix** is the flagged box of x_0 (opened before line 3, closed at line 11); **obtain** is line 3 (∀E at x_0) and the ∃E box of y_0 (lines 4–10), with lines 5–7 using the second premise on the witness; **choose** is line 9, ∃I with the witness y_0.

Exercise. B3

An integer n is **odd** when $n = 2k+1$ for some integer k. Over $\mathbb{Z}$, derive

$$\exists k\, (m = 2 \cdot k + 1),\ \exists k\, (n = 2 \cdot k + 1) \ \vdash\ \exists k\, (m \cdot n = 2 \cdot k + 1),$$

with one line justified "arith". Then write the proof as three sentences of prose.

Solution. Idea. As in Example (The sum of two even integers is even): open both boxes early, with different names, and postpone the goal until the arithmetic shows the witness.

1		$\exists k\,(m = 2 \cdot k + 1)$	
2		$\exists k\,(n = 2 \cdot k + 1)$	
3	k_0	$m = 2 \cdot k_0 + 1$	asm, for $\exists$E
4	l_0	$n = 2 \cdot l_0 + 1$	asm, for $\exists$E
5		$m \cdot n = 2 \cdot (2 \cdot k_0 \cdot l_0 + k_0 + l_0) + 1$	arith 3, 4
6		$\exists k\,(m \cdot n = 2 \cdot k + 1)$	$\exists$I 5
7		$\exists k\,(m \cdot n = 2 \cdot k + 1)$	$\exists$E 2, 4–6
8		$\exists k\,(m \cdot n = 2 \cdot k + 1)$	$\exists$E 1, 3–7

Line 5: $(2k_0 + 1)(2l_0 + 1) = 4k_0l_0 + 2k_0 + 2l_0 + 1 = 2(2k_0l_0 + k_0 + l_0) + 1$. Line 6 is $\exists$I with $\varphi = (m \cdot n = 2 \cdot k + 1)$ and the witness $2 \cdot k_0 \cdot l_0 + k_0 + l_0$. At line 7, l_0 is in none of line 6, line 2, the premises, and the open assumption on line 3; at line 8, k_0 is in none of line 7, line 1, the premises, and no box is open.

In prose: *Choose integers k and l with $m = 2k + 1$ and $n = 2l + 1$. Then $mn = 4kl + 2k + 2l + 1 = 2(2kl + k + l) + 1$. Since $2kl + k + l$ is an integer, mn is odd.*

Exercise. B4: Derive or break

For each claim, derive it with the rules so far, or break it with a counter-reading.

(a) $\exists x\,\forall y\,R(x, y) \vdash \exists x\,R(x, x)$.
(b) $\forall x\,\exists y\,R(x, y) \vdash \exists x\,R(x, x)$.
(c) $\forall x\,(P(x) \to Q(x)),\ \exists x\,\neg Q(x) \vdash \exists x\,\neg P(x)$.
(d) $\forall x\,(P(x) \to Q(x)),\ \exists x\,\neg P(x) \vdash \exists x\,\neg Q(x)$.
(e) $\exists x\,P(f(x)) \vdash \exists y\,P(y)$.
(f) $\exists y\,P(y) \vdash \exists x\,P(f(x))$.
(g) $\forall x\,(P(x) \to \exists y\,R(x, y)),\ \exists x\,P(x) \vdash \exists y\,\exists x\,R(x, y)$.

Solution.

(a) Derive. Obtain the witness, and instantiate its universal at the witness itself.

1		$\exists x\,\forall y\,R(x, y)$	
2	x_0	$\forall y\,R(x_0, y)$	asm, for $\exists$E
3		$R(x_0, x_0)$	$\forall$E 2
4		$\exists x\,R(x, x)$	$\exists$I 3
5		$\exists x\,R(x, x)$	$\exists$E 1, 2–4

(b) Break. The domain $\{0, 1\}$, with R true of the pairs $(0, 1)$ and $(1, 0)$ only. Each element is related to something, and no element to itself.

(c) Derive. The $\exists$E box opens first, aimed at the goal; inside it, a $\neg$I box.

1	$\forall x\,(P(x) \to Q(x))$	
2	$\exists x\,\neg Q(x)$	
3	x_0 | $\neg Q(x_0)$	asm, for $\exists$E
4	$P(x_0) \to Q(x_0)$	$\forall$E 1
5	$P(x_0)$	asm, for $\neg$I
6	$Q(x_0)$	$\to$E 4, 5
7	$\bot$	$\neg$E 6, 3
8	$\neg P(x_0)$	$\neg$I 5–7
9	$\exists x\,\neg P(x)$	$\exists$I 8
10	$\exists x\,\neg P(x)$	$\exists$E 2, 3–9

(d) Break. The domain $\{0, 1\}$, with P true of nothing and Q true of both elements. The first premise holds (no hypothesis $P(x)$ is true), the second holds since $\neg P(0)$, and no element lacks Q.

(e) Derive. The witness of the conclusion is the compound term $f(x_0)$.

1	$\exists x\,P(f(x))$	
2	x_0 | $P(f(x_0))$	asm, for $\exists$E
3	$\exists y\,P(y)$	$\exists$I 2
4	$\exists y\,P(y)$	$\exists$E 1, 2–3

(f) Break. The domain $\{0, 1\}$, with f sending both elements to 0, and P true of 1 only. Some element has P, and no value of f does.

(g) Derive. Two boxes with different names, the second inside the first.

1	$\forall x\,(P(x) \to \exists y\,R(x, y))$	
2	$\exists x\,P(x)$	
3	x_0 | $P(x_0)$	asm, for $\exists$E
4	$P(x_0) \to \exists y\,R(x_0, y)$	$\forall$E 1
5	$\exists y\,R(x_0, y)$	$\to$E 4, 3
6	y_0 | $R(x_0, y_0)$	asm, for $\exists$E
7	$\exists x\,R(x, y_0)$	$\exists$I 6
8	$\exists y\,\exists x\,R(x, y)$	$\exists$I 7
9	$\exists y\,\exists x\,R(x, y)$	$\exists$E 5, 6–8
10	$\exists y\,\exists x\,R(x, y)$	$\exists$E 2, 3–9

At line 9 the open assumption $P(x_0)$ does not contain y_0; at line 10 no box is open.

The pairs: (c) and (d) are the contrapositive and its look-alike (denying the antecedent); (e) and (f) are converses; (a) and (b) differ in the order of the quantifiers. Worked examples of these types: Proposition 5.6.1 and its Quick check (a counter-reading); Example (Derivations

with existential elimination) (b); Example (The sum of two even integers is even); the two-dot counter-reading of the converse of Proposition 5.6.2.

Marking scheme. (10 points)

- 1 each: (a), (b), (d), (e), (f), with the derivation or the counter-reading in full.
- 2: (c), with the ∃E box opened first and aimed at the goal, and $\neg P(x_0)$ obtained inside it (by a ¬I box or by MT).
- 2: (g), with two boxes, different names, closed in the right order.
- 1: the closing pairing.

Fatal errors (0 for the claim): an ∃E box closed on a line that mentions its witness; a counter-reading in which a premise is false.

5.6.10.3 C. Going deeper

Exercise. C1

Derive

$$\exists x\,\forall y\,R(x,y),\ \forall x\,\forall y\,(R(x,y)\to R(y,x))\ \vdash\ \forall y\,\exists x\,R(y,x),$$

and check the side condition of every ∃E and ∀I line. Then give a reading over $\mathbb{Z}$ in which both premises are true, and say which integer the derivation's witness is.

Solution (Hint 1). The shape is the one of Proposition 5.6.2: fix, obtain, choose. Which line does symmetry have to turn round?

Solution (Hint 2). Open a flagged box for y_0, and inside it the ∃E box of the first premise, with the assumption $\forall y\,R(x_0,y)$.

Solution. Idea. The goal is $\forall y$: fix y_0. The new goal $\exists x\,R(y_0,x)$ is postponed. The first premise is an existential: obtain x_0 with $\forall y\,R(x_0,y)$, and instantiate it at the term the goal mentions, y_0: $R(x_0,y_0)$. The goal needs $R(y_0,t)$, with y_0 first, so the second premise turns the line round, instantiated at x_0 and y_0 in that order (read off the line to be turned). Then choose the witness x_0.

1			$\exists x\, \forall y\, R(x, y)$	
2			$\forall x\, \forall y\, (R(x, y) \to R(y, x))$	
3	y_0	x_0	$\forall y\, R(x_0, y)$	asm, for $\exists$E
4			$R(x_0, y_0)$	$\forall$E 3
5			$\forall y\, (R(x_0, y) \to R(y, x_0))$	$\forall$E 2
6			$R(x_0, y_0) \to R(y_0, x_0)$	$\forall$E 5
7			$R(y_0, x_0)$	$\to$E 6, 4
8			$\exists x\, R(y_0, x)$	$\exists$I 7
9			$\exists x\, R(y_0, x)$	$\exists$E 1, 3–8
10			$\forall y\, \exists x\, R(y, x)$	$\forall$I 3–9

Line 9 ($\exists$E): x_0 does not appear in $\chi = \exists x\, R(y_0, x)$, nor in the existential of line 1, nor in either premise; the only box open at line 9 is the flagged box of y_0, which has no assumption. Line 8 is a line of the $\exists$E box itself.

Line 10 ($\forall$I): the flagged box immediately contains the $\exists$E box, so its range starts at line 3. Its last line, line 9, is a line of the box itself and is $(\exists x\, R(y, x))[y_0/y]$. The name y_0 appears in neither premise and not in line 10, and no box is open.

A reading. Over $\mathbb{Z}$, let $R(x, y)$ mean "$x \cdot y = 0$". The second premise holds since $x \cdot y = y \cdot x$. The first holds with the witness 0, and only with it: for $x \neq 0$, the instance $y = 1$ gives $x \cdot 1 = x \neq 0$. The conclusion says that every integer has a partner with product 0, which is true. The derivation's x_0 names the witness of the first premise, so under this reading it can only be 0, and the derivation gives every y the partner 0; the derivation itself never uses which integer it is.

Marking scheme. (10 points) - 2: opens a flagged box for the goal's $\forall y$ with a fresh name, and postpones the existential. - 2: opens the $\exists$E box of the first premise with a **fresh** witness, aimed at the current goal. - 2: instantiates the assumption at y_0 and the symmetry premise at x_0, y_0 in that order, and fires it. - 2: $\exists$I with the witness x_0, inside the $\exists$E box. - 1: checks the side conditions of the $\exists$E and $\forall$I lines clause by clause. - 1: a reading over $\mathbb{Z}$ in which both premises hold, with the witness identified (or, if the reading does not determine it, correctly described as whichever witness the first premise has). Fatal errors (0 for the step): an $\exists$E box closed on a line containing its witness; $\forall$I over a name in an open assumption; the symmetry premise instantiated in the wrong order and then used as if it fired.

Exercise. C2: Find the error, used later

Here is a "derivation" of $\exists x\, (P(x) \wedge Q(x)) \vdash \exists x\, P(x)$.

1	$\exists x\,(P(x) \wedge Q(x))$	
2	x_0 | $P(x_0) \wedge Q(x_0)$	asm, for $\exists$E
3	$P(x_0)$	$\wedge$E 2
4	$P(x_0)$	$\exists$E 1, 2–3
5	$\exists x\,P(x)$	$\exists$I 4

(a) Is the conclusion derivable from the premise? Justify your answer.

(b) Does this derivation prove it? If not, name the exact line that fails, the rule, and the clause of its side condition that is broken.

(c) Show that the illegal step itself is unsound: use the same step to "derive", from a premise true over $\mathbb{Z}$, a sentence false over $\mathbb{Z}$.

Solution (Hint 1). For (a), the conclusion is half of the conclusion of a result proved in this section.

Solution (Hint 2). For (c), drop Q: start from the premise $\exists x\,P(x)$ and take the same illegal step as line 4.

Solution. *(a)* Yes, the conclusion is derivable: by Proposition 5.6.1 the premise gives $\exists x\,P(x) \wedge \exists x\,Q(x)$, and $\wedge$E gives $\exists x\,P(x)$. Directly, moving the $\exists$I inside the box repairs the derivation:

1	$\exists x\,(P(x) \wedge Q(x))$	
2	x_0 | $P(x_0) \wedge Q(x_0)$	asm, for $\exists$E
3	$P(x_0)$	$\wedge$E 2
4	$\exists x\,P(x)$	$\exists$I 3
5	$\exists x\,P(x)$	$\exists$E 1, 2–4

At line 5, x_0 is not in line 4, not in line 1, and no box is open.

(b) No. Line 4 fails: it is $\exists$E with $\chi = P(x_0)$, and x_0 appears in χ, which breaks the clause "x_0 does not appear in χ" of Definition 5.6.1. The witness has escaped its box. Lines 1–3 are correct, and line 5 would be a correct $\exists$I if line 4 were a legitimate line. A true conclusion does not make the argument correct: the conclusion is derivable, but not by this derivation.

(c) The same step, from $\exists x\,P(x)$ alone:

1	$\exists x\, P(x)$	
2	$x_0 \mid P(x_0)$	asm, for $\exists$E
3	$P(x_0)$	$\exists$E 1, 2–2
4	$x_0 \mid P(x_0)$	R 3
5	$\forall x\, P(x)$	$\forall$I 4–4

Line 3 is the illegal step of line 4 above. Line 4 repeats line 3 into a flagged box, and line 5 is a correct $\forall$I: x_0 appears in no premise, in no open assumption, and not in line 5. Over $\mathbb{Z}$ with P as "even", the premise "some integer is even" is true and the conclusion "every integer is even" is false. So the step from $\exists x\, P(x)$ to $P(x_0)$ can lead from truth to falsehood; it is a counterexample to the **step**, while the conclusion of the original sequent has no counterexample at all. (This is the same failure as the derivation after Definition 5.6.1 that "proves" every natural number exceeds 5.)

Marking scheme. (10 points) - 2: (a) says the conclusion is derivable and gives a correct derivation or cites Proposition 5.6.1 with $\wedge$E. - 3: (b) names line 4, the rule $\exists$E, and the clause "x_0 not in χ". - 1: (b) says that the other lines are correct, so the error is in that one line. - 4: (c) gives a derivation that uses the same illegal step and is otherwise legal, with a reading in which the premise is true and the conclusion false. Fatal errors (0 for the step): answering (a) "no" because the derivation is wrong; giving a counterexample to the conclusion of the original sequent (there is none); blaming line 5.

Exercise. C3: The diagonal of a list

Let $A_0, A_1, A_2, \ldots$ be sets of natural numbers, one for each $n \in \mathbb{N}$, and let $D = \{n \in \mathbb{N} : n \notin A_n\}$.

(a) Over the domain $\mathbb{N}$, read $R(x, y)$ as "$x \in A_y$". Follow the proof of Theorem 5.6.3 in this reading to show that D is none of the sets A_m.

(b) Compute D for the list $A_n = \{k \in \mathbb{N} : k > n\}$, and for the list $A_n = \{k \in \mathbb{N} : k \text{ is a multiple of } n+1\}$. In each case, find a number at which D differs from A_0, from A_1 and from A_2.

Solution (Hint). In (a), the existential of line 1 is "some m has $D = A_m$". What does the definition of D say about every n, when $D = A_m$?

Solution. (a) Suppose $D = A_m$ for some m: this is the existential of line 1, with its witness named m (line 2). By the definition of D, for every n: $n \in A_m$ exactly when $n \notin A_n$; that is $\forall x\,(R(x, m) \leftrightarrow \neg R(x, x))$. At $n = m$ (line 3): $m \in A_m$ exactly when $m \notin A_m$, which is impossible (lines 4–9). So no m has $D = A_m$.

(b) First list: $n > n$ is false, so $n \notin A_n$ for every n, and $D = \mathbb{N}$. D differs from A_m at m: $m \in D$ and $m \notin A_m$ ($0 \notin A_0$, $1 \notin A_1$, $2 \notin A_2$).

Second list: n is a multiple of $n + 1$ only for $n = 0$ ($0 = 0 \cdot 1$; for $n \geq 1$, $0 < n < n + 1$), so $D = \{1, 2, 3, \ldots\}$. $A_0 = \mathbb{N}$ contains $0 \notin D$; $A_1 = \{0, 2, 4, \ldots\}$ lacks $1 \in D$; $A_2 = \{0, 3, 6, \ldots\}$

lacks $2 \in D$.

In both, D and A_m differ at the number m itself, as the proof says they must.

Marking scheme. (10 points)

- 2: the assumption $D = A_m$ set up as the witness of an existential.
- 2: the universal $\forall x\,(R(x,m) \leftrightarrow \neg R(x,x))$ read off the definition of D.
- 2: the instance at m and the contradiction.
- 2: the first list's D, with the three differences.
- 2: the second list's D (the case $n = 0$ justified), with the three differences.

Fatal error (0 for (a)): instantiating at some number other than m.

5.7 Equality

Sections 5.3 to 5.6 gave the four quantifier rules, and with them every logical symbol of the language of Section 5.2 (the connectives, $\bot$, the quantifiers, and equality) has rules, except one. So far the equals sign has been treated like a predicate symbol: $a = b$ was an atomic formula like $P(a)$, and no rule looked inside it. Yet almost every calculation in mathematics is a chain of equations, and every use of "the" (the limit, the supremum, the inverse) quietly asserts that something is unique. This section adds the two rules for equality, =I and =E, which say together that *equals may be replaced by equals*. Section 5.8 applies them to the phrase "there is exactly one", which is an existential statement together with an equation. This section ends by completing the table of rules, phrases, and Lean tactics that Chapter 1 began.

5.7.1 What the rules so far cannot do with an equation

Here is the first line of a calculation any reader has written:

> Since $a = b$ and $a + a = 2a$, we have $b + b = 2b$.

Nothing could be simpler, and nothing in Sections 5.3–5.6 derives it: so far the rules treat $=$ as they treat any two-place predicate symbol, and from $R(a,b)$ and $P(a)$ no rule gives $P(b)$. Reading $R(a,b)$ as $a < b$ and $P(x)$ as $x < 1$ over $\mathbb{Z}$, with $a = 0$ and $b = 5$, both premises are true and $P(b)$ is false.

What makes the step correct is what $=$ means: a and b are names of **the same object**. Any statement true of the object is true of it whichever name is used. Two things follow, and they are the two rules of this section.

1. Every term names the same object as itself. Nothing needs to be assumed to write $t = t$.
2. If t_1 and t_2 name the same object, then a statement about t_1 may be rewritten as the same statement about t_2.

The first rule is unusual: it cites no earlier line at all. Chapter 1 anticipated this when it defined rules (Section 1.2): the definition of a rule schema (Definition 1.2.1) was worded so that a rule of Chapter 5 could have no given lines at all. This is that rule.

5.7.2 The rules for equality

Everything is equal to itself, and equals may be replaced by equals.

Definition 5.7.1. Rules for Equality

$$\begin{array}{lll|ll} & & & i & t_1 = t_2 & \\ k & t = t & =\text{I} \qquad\qquad & j & \varphi[t_1/x] & \\ & & & k & \varphi[t_2/x] & =\text{E } i,\ j \end{array}$$

The rule =I (**equality introduction**, or reflexivity) writes $t = t$ for **any** term t, at any line, citing nothing. The rule =E (**equality elimination**, or substitution of equals) cites an equation $t_1 = t_2$ on line i and a line j of the form $\varphi[t_1/x]$, both available at line k, and writes $\varphi[t_2/x]$ on line k. Here φ is **any** formula and x any variable, subject to one side condition: **both t_1 and t_2 are free for x in φ** (Definition 5.2.6).

In words: =I says that every term equals itself. =E says that when $t_1 = t_2$ is known, a line that says something about t_1 may be rewritten as the same thing said about t_2. The formula φ is the "something": a statement with a hole x, which line j fills with t_1 and line k fills with t_2.

Mark the holes first. To use =E with $t_1 = t_2$, write down the line you have and draw a box round each occurrence of t_1 you want to replace. The boxes are the holes: φ is the line with a variable in each box, and the new line puts t_2 in every box. With $a = b$:

$$\boxed{a} + \boxed{a} = 2 \cdot \boxed{a} \text{ gives } b + b = 2 \cdot b, \qquad \boxed{a} + a = 2 \cdot a \text{ gives } b + a = 2 \cdot a.$$

Occurrences without a box stay as they are.

Read the definition clause by clause.

1. =I **cites nothing.**
   - *Schema:* the left block above, $t = t$, citing nothing.
   - *Reading:* every term equals itself, so the line may stand anywhere, even first in a derivation with no premises.
   - *Example:* "$a + b = a + b$ — =I", the first line of every chain below.
2. **The equation is read from left to right.** The cited equation $t_1 = t_2$ licenses replacing t_1 by t_2, and not the other way round. From $a = b$ and $P(b)$, the rule does **not** give $P(a)$, because $P(b)$ is $\varphi[b/x]$ for $\varphi = P(x)$, and the equation would have to be $b = a$. The reverse direction needs one more step, symmetry, which we derive below.
3. **Some or all occurrences.** The choice of φ decides which occurrences of t_1 are replaced: exactly those that φ marks with x (the boxes above). Occurrences of t_1 that φ writes out are left alone. Choosing φ is the whole art of using the rule.
4. **Free for x.** Neither term may have a variable captured by a quantifier of φ, the condition of Definition 5.2.6 that the rules of Sections 5.3 and 5.5 also carry. The nonexample below shows what goes wrong without it.

One line, many readings. Take the equation $a = b$ and the line $a + a = 2 \cdot a$. The term a occurs three times in the line, and φ may mark any of them. Some of the choices:

φ	$\varphi[a/x]$ (the cited line)	$\varphi[b/x]$ (the new line)
$x + a = 2 \cdot a$	$a + a = 2 \cdot a$	$b + a = 2 \cdot a$
$a + x = 2 \cdot a$	$a + a = 2 \cdot a$	$a + b = 2 \cdot a$
$a + a = 2 \cdot x$	$a + a = 2 \cdot a$	$a + a = 2 \cdot b$
$x + x = 2 \cdot x$	$a + a = 2 \cdot a$	$b + b = 2 \cdot b$

Each of the three occurrences is marked or not, so there are $2^3 = 8$ choices in all, and each is a legal step. The fourth row is the step of the opening calculation.

↻ **Second reading** · *Degenerate choices*

One more row belongs to the table:

φ	$\varphi[a/x]$ (the cited line)	$\varphi[b/x]$ (the new line)
$a + a = 2 \cdot a$	$a + a = 2 \cdot a$	$a + a = 2 \cdot a$

This row is the **degenerate** choice: φ contains no x, so $\varphi[a/x]$ and $\varphi[b/x]$ are both φ, and the rule copies the line unchanged. It is legal and useless, and it shows that "some occurrences" includes "none". Another degenerate case: the equation may be $t = t$ itself, and then =E again returns the cited line.

Quick check. From the equation $x_0 = 3$ and the line $x_0 \cdot x_0 > x_0$, which formula φ makes =E give $3 \cdot 3 > x_0$? Which gives $x_0 \cdot 3 > 3$?

Solution. For $3 \cdot 3 > x_0$, take $\varphi = (z \cdot z > x_0)$, with a variable z that appears nowhere else: $\varphi[x_0/z]$ is the cited line and $\varphi[3/z]$ is $3 \cdot 3 > x_0$. For $x_0 \cdot 3 > 3$, take $\varphi = (x_0 \cdot z > z)$. The hole has to be a variable other than x_0, since the occurrences of x_0 that stay must be written out in φ.

Quick check. From $a = b$ and the line $a \cdot a = a + 1$, which occurrences do you box to get $a \cdot b = b + 1$?

Solution. The second and the third a: $\varphi = (a \cdot z = z + 1)$, and $\varphi[a/z]$ is the given line, $\varphi[b/z]$ the new one.

A nonexample: capture. Take the equation $a = y$, with y a variable, and the line $\exists y\,(y < a)$. It is tempting to "substitute y for a" and write $\exists y\,(y < y)$. With $\varphi = \exists y\,(y < x)$, the cited line is $\varphi[a/x]$, and the new line would be $\varphi[y/x]$. But y is **not** free for x in φ: the x sits inside $\exists y$,

and the y put there would be captured. The side condition fails, and rightly. Read over $\mathbb{Z}$, with a and y both naming 0, the premises say "$a = y$" (true) and "some integer is less than 0" (true), while $\exists y\,(y < y)$ says that some integer is less than itself (false).

1	$a = y$	
2	$\exists y\,(y < a)$	
3	$\exists y\,(y < y)$	=E 1, 2

Line 3 breaks the side condition of Definition 5.7.1. What the argument meant to say needs a different name for the bound variable: $\exists z\,(z < a)$, which is the same sentence up to renaming (Section 5.2), becomes $\exists z\,(z < y)$, which is true when y names 0.

Why these two rules. They are exactly the two facts about "same object" listed above, and nothing else about equality needs to be assumed. Symmetry ($s = t$ gives $t = s$), transitivity ($s = t$ and $t = u$ give $s = u$), and "applying the same function to both sides" are all derived below. So the two rules play for $=$ the part that the introduction and elimination rules played for each connective: one builds an equation, one uses it.

Warning. **A bound variable is not a place to substitute.** From $x = 0$ and $\forall x\,(x = x)$, the line $\forall x\,(0 = x)$ does **not** follow: the x's of $\forall x\,(x = x)$ are bound, and =E replaces only free occurrences. Over $\mathbb{Z}$, with x naming 0, the premises are true and $\forall x\,(0 = x)$ is false.

Second reading · *Why no choice of the hole works*

To write $\forall x\,(x = x)$ as $\varphi[x/z]$, either z does not occur free in φ, and then $\varphi[0/z]$ is the unchanged line, or z occurs free in φ (for instance, φ is $\forall x\,(z = x)$), and then every free z lies inside $\forall x$, so x is not free for z in φ and the side condition fails.

5.7.3 Symmetry, transitivity and congruence

Everything we expect of equality now follows from the two rules. Each derivation below is short, and the only thought in each is the choice of φ.

It turns out that the equation in =E may be read in either direction after all, at the cost of two extra lines (an =I and an =E; after the abbreviation (sym) below, just one), and that equations chain:

Proposition 5.7.1. Equality Is Symmetric and Transitive

For all terms s, t, u:

(a) $s = t \vdash t = s$;

(b) $s = t,\ t = u \vdash s = u$.

Idea. For (a), the goal $t = s$ must be the last line of an =E that uses the premise $s = t$ (it is the only equation available), so it has the form $\varphi[t/x]$, and the cited line $\varphi[s/x]$ must already be on a line. Try $\varphi = (x = s)$: then $\varphi[t/x]$ is $t = s$, the goal, and $\varphi[s/x]$ is $s = s$, which =I supplies for free. For (b), rewrite the t of the first premise using the second: $\varphi = (s = x)$.

Proof.

(a) With $\varphi \coloneqq (x = s)$, for a variable x that occurs in neither s nor t, the line $s = s$ is $\varphi[s/x]$ and the goal $t = s$ is $\varphi[t/x]$. Both terms are free for x in φ, since φ has no quantifier. So the following is a derivation:

1	$s = t$	
2	$s = s$	=I
3	$t = s$	=E 1, 2

(b) With $\varphi \coloneqq (s = x)$, for a variable x that does not occur in s, the first premise $s = t$ is $\varphi[t/x]$ and the goal $s = u$ is $\varphi[u/x]$; again φ has no quantifier. So the following is a derivation:

1	$s = t$	
2	$t = u$	
3	$s = u$	=E 2, 1

Line 3 of (b) cites the equation $t = u$ on line 2 as the rule's line i, and $s = t$ on line 1 as its line j. This proves (a) and (b).

■

Notice where the "reflexivity" line went in (a): the equation $s = t$ was not rewritten at all. It was the **tool**, and the line it rewrote, $s = s$, was manufactured for the purpose. This is the typical use of =I: it supplies a line for =E to work on.

From now on, symmetry and transitivity may be used as one line, with the justifications "(sym)" and "(trans)": each abbreviates the three-line derivation above with the given terms in place of s, t, u. This is the license of Definition 2.3.1, extended from statement letters to terms. The extension is harmless for the same reason: the derivations above used nothing about s, t, u except that they are terms, so they can be pasted in with any terms.

The next derivation is the rule behind "apply f to both sides". Let f be a one-place function symbol, and a, b constants.

1	$a = b$	
2	$f(a) = f(a)$	=I
3	$f(a) = f(b)$	=E 1, 2

Here $\varphi = (f(a) = f(x))$: it keeps the first a and marks the second. With $\varphi = (f(x) = f(x))$ instead, the same two lines would give $f(b) = f(b)$, which is true and useless. Nothing is special about f: the same three lines give $a^2 = b^2$ from $a = b$, or $a + c = b + c$, or $|a| = |b|$. Every "do

the same thing to both sides of an equation" is this derivation.

The last result of this group has no premises. It says that equals share every property, and it is called **Leibniz's Law**.

Proposition 5.7.2. Leibniz's Law

For every one-place predicate symbol P, $\vdash \forall x\, \forall y\, (x = y \to (P(x) \to P(y)))$.

In words: whatever is true of x is true of anything equal to it. The rule =E is this law for every formula at once; the seven-line derivation is in the fold below.

Second reading · *Leibniz's Law, derived*

Proof. Finding it. The goal is $\forall x$, so the search method of Section 3.1, with the step of Section 5.4, opens a flagged box for x_0; the new goal $\forall y\,(\dots)$ opens a second flagged box, for y_0. The goal $x_0 = y_0 \to (P(x_0) \to P(y_0))$ is an implication, so a box opens with $x_0 = y_0$, and the new goal $P(x_0) \to P(y_0)$ opens another, with $P(x_0)$. Now the goal is $P(y_0)$, the resources are $x_0 = y_0$ and $P(x_0)$, and one =E with $\varphi = P(x)$ finishes. Every line but one was written by a goal-directed step.

1	x_0	y_0	$x_0 = y_0$	asm, for $\to$I
2			$P(x_0)$	asm, for $\to$I
3			$P(y_0)$	=E 1, 2
4			$P(x_0) \to P(y_0)$	$\to$I 2–3
5			$x_0 = y_0 \to (P(x_0) \to P(y_0))$	$\to$I 1–4
6			$\forall y\,(x_0 = y \to (P(x_0) \to P(y)))$	$\forall$I 1–5
7			$\forall x\, \forall y\, (x = y \to (P(x) \to P(y)))$	$\forall$I 1–6

The two $\forall$I lines cite the flagged boxes as 1–5 and 1–6: each flagged box immediately contains another box, so its range starts at that box's first line (the convention of Section 5.4). At line 6 the name y_0 appears in no premise (there are none), in no open assumption (the boxes on lines 1 and 2 have closed), and not in the conclusion; at line 7 the same holds for x_0. The derivation has 7 lines.

■

The claim is one of the classical "axioms of equality", often taken as a starting point in place of =E. Here it is a consequence. Conversely, with the claim as a premise, $\forall$E twice and $\to$E twice would do the work of =E for the formula $P(x)$; the rule does it for every formula at once.

Warning. **Words about knowledge are not formulas.** =E rewrites formulas of the language, and "it is obvious that …", "the pupil checked that …", "it is known that …"

are not among its connectives. A pupil checked that $7 \cdot 13 = 91$, and indeed $7 \cdot 13 = 91$. Replacing $7 \cdot 13$ by 91 inside the report gives "the pupil checked that $91 = 91$", a different claim about what the pupil did. Inside a formula every free occurrence of a term may be rewritten, subject to the side condition of Definition 5.7.1; inside words about what someone knows or checked, none may.

5.7.3.1 Equal sets have the same elements

Example. Equal Sets Have the Same Elements

Read the domain as consisting of sets, with $\in$ a two-place predicate symbol written between its terms (shorthand rule 1 of Section 5.2). Show that $\vdash \forall x \forall y \, (x = y \to \forall z \, (z \in x \leftrightarrow z \in y))$.

Solution.

1	x_0	y_0		$x_0 = y_0$	asm, for $\to$I
2				z_0 | $z_0 \in x_0$	asm, for $\leftrightarrow$I
3				$z_0 \in x_0$	asm, for $\leftrightarrow$I
4				$z_0 \in x_0 \leftrightarrow z_0 \in x_0$	$\leftrightarrow$I 2–2, 3–3
5				$z_0 \in x_0 \leftrightarrow z_0 \in y_0$	=E 1, 4
6				$\forall z \, (z \in x_0 \leftrightarrow z \in y_0)$	$\forall$I 2–5
7			$x_0 = y_0 \to \forall z \, (z \in x_0 \leftrightarrow z \in y_0)$		$\to$I 1–6
8		$\forall y \, (x_0 = y \to \forall z \, (z \in x_0 \leftrightarrow z \in y))$			$\forall$I 1–7
9	$\forall x \forall y \, (x = y \to \forall z \, (z \in x \leftrightarrow z \in y))$				$\forall$I 1–8

The two boxes of line 4 are one-line boxes, as $P \to P$ is shown, and each flagged box immediately contains another box and is cited from its first line. Line 5 is =E with $\varphi = (z_0 \in x_0 \leftrightarrow z_0 \in w)$: line 4 is $\varphi[x_0/w]$, and the equation of line 1 refills the hole with y_0; as in the proof of symmetry, $\leftrightarrow$I manufactured a line for =E to rewrite.

The converse, $\forall z \, (z \in x \leftrightarrow z \in y) \to x = y$, breaks. Here is a counter-reading to it: two dots a and b, with nothing an element of anything. With x naming a and y naming b, both sides of $z \in x \leftrightarrow z \in y$ are false for every z, so the hypothesis is true; and $a = b$ is false. For sets, the converse is taken as an axiom, the axiom of extensionality: a set is determined by its elements.

5.7.4 Calculations are chains of substitutions

A calculation in a proof looks like this:

$$a + b = 2 \cdot b + b = 2 \cdot 3 + 3 = 9,$$

written under the premises $a = 2 \cdot b$ and $b = 3$. The chain is a single statement to the eye, but logically it is several equations, each obtained from the last. Here is the derivation it abbreviates, with one line of arithmetic, justified "arith" as in Section 5.3.

A chain as a derivation. From $b = 3$ and $a = 2 \cdot b$, derive $a + b = 9$.

1	$b = 3$	
2	$a = 2 \cdot b$	
3	$a + b = a + b$	=I
4	$a + b = 2 \cdot b + b$	=E 2, 3
5	$a + b = 2 \cdot 3 + 3$	=E 1, 4
6	$2 \cdot 3 + 3 = 9$	arith
7	$a + b = 9$	=E 6, 5

The choices of φ, with z a variable not occurring in the terms:

- line 4: $\varphi = (a + b = z + b)$, rewriting the a on the **right** of line 3 and keeping the one on the left;
- line 5: $\varphi = (a + b = 2 \cdot z + z)$, rewriting **both** remaining b's on the right at once;
- line 7: $\varphi = (a + b = z)$, rewriting the whole right-hand side by the equation of line 6.

The pattern is general. A chain $t_1 = t_2 = \cdots = t_n$ starts with $t_1 = t_1$ by =I, keeps t_1 fixed on the left, and rewrites the right-hand side once per link, each time by =E with the equation that justifies the link. Each link is either a premise, as on lines 4 and 5, or an arithmetic fact, as on line 7. The last line is $t_1 = t_n$. Transitivity (Proposition 5.7.1 (b)) is the same move made once; a chain makes it $n - 1$ times.

A **mixed chain**, such as $a^2 + 1 = 10 \leq 10 + b^2$ with $a = 3$, is not made of =E alone. Its $=$ links are =E steps as above, and a $\leq$ link needs a fact about $\leq$: here that $b^2 \geq 0$, so that adding b^2 does not decrease a number. The $=$ link before it then carries over by one more =E, with $\varphi = (z \leq 10 + b^2)$ and the equation $a^2 + 1 = 10$ turned round by symmetry (Proposition 5.7.1 (a)); transitivity of $\leq$ is needed only where two $\leq$ links meet, as in $x \leq y \leq z$. Those are facts about order, not about equality. They enter derivations as "arith" lines, one per link.

Quick check. Line 5 of the chain rewrites both b's on the right of line 4 at once. Which φ would rewrite only the b inside $2 \cdot b$, and what line would it give? How would the chain then reach line 5?

Solution. Take $\varphi = (a + b = 2 \cdot z + b)$. Then line 4 is $\varphi[b/z]$, and =E 1, 4 gives $\varphi[3/z]$, the line $a + b = 2 \cdot 3 + b$. One more =E with the same equation $b = 3$ and $\varphi = (a + b = 2 \cdot 3 + z)$ then gives $a + b = 2 \cdot 3 + 3$, which is line 5. Marking both occurrences at once saves that line.

Quick check. Under the premise $a = 2$, the chain $a + 1 = 2 + 1 = 3 \leq 4$ has three links. Which is an =E step, and which are "arith" lines?

Solution. The first link is =E: from $a + 1 = a + 1$ (=I) and $a = 2$, with $\varphi = (a + 1 = z + 1)$, it writes $a + 1 = 2 + 1$. The links $2 + 1 = 3$ and $3 \leq 4$ are facts of school arithmetic, written as "arith" lines, and each is then carried into the chain by one more =E (with $\varphi = (a + 1 = z)$,

then with $\varphi = (z \leq 4)$ and the equation turned round); the last link is about order, not equality.

5.7.5 What equality sounds like in a proof

In a written proof, =E is the word "substituting", or no word at all: the chain $a+b = 2b+b = 9$ is read as a sequence of substitutions, and the reader supplies the φ of each link. "By symmetry" or "rewriting the equation as $t = s$" is Proposition 5.7.1 (a). "Applying f to both sides", "squaring both sides", "adding c to both sides" are the congruence derivation above. And =I is almost never written, because $t = t$ goes without saying; it is visible only in the first line of a formal chain, and in a proof assistant, where `rfl` proves $t = t$.

Here is a proof from school algebra, read link by link.

> **Claim.** If $a + b = 0$ and $a + c = 0$, then $b = c$.
>
> *Proof.* $b = b + 0 = b + (a + c) = (b + a) + c = (a + b) + c = 0 + c = c$.

The first link is the fact $y + 0 = y$ for every y, applied at $y = b$ (an instance by ∀E) and read from right to left: the chain needs $b = b + 0$, which is the instance $b + 0 = b$ turned round by symmetry. The second link rewrites the 0 by the premise $a + c = 0$, again turned round, since the chain needs 0 to become $a + c$. The third and fourth are instances of associativity and commutativity. The fifth rewrites $a + b$ by the premise $a + b = 0$ in the direction written. The last is the instance $0 + c = c$. So the claim says that a number has **at most one** additive inverse; the facts it uses are the field axioms, here school algebra.

5.7.6 The search method: one more step

Each of Sections 5.3 to 5.6 added a step to the search method of Section 3.1. Equality adds one.

- **An equation.** When an equation $t_1 = t_2$ is available and the goal differs from an available line only in some occurrences of t_1 and t_2, look for the φ that makes one =E turn the line into the goal: mark with x exactly the places where they differ. If the equation points the wrong way, turn it round first (sym). If no available line will do, make one with =I, as the proof of Proposition 5.7.1 did.

In the Givens/Goal table, =E turns a Given into a new Given with some holes refilled, and =I adds $t = t$ to the Givens at no cost.

Example. Instances of =E

In each row the hole is z; the cited line fills it with the left side of the equation, the new line with the right side.

Equation $t_1 = t_2$	φ	Cited line $\varphi[t_1/z]$	New line $\varphi[t_2/z]$
$f(c) = g(c) + 1$	$P(z) \wedge R(f(c), c)$	$P(f(c)) \wedge R(f(c), c)$	$P(g(c) + 1) \wedge R(f(c), c)$
$\varepsilon/2 = \delta$	$\lvert x - a\rvert < z \to \lvert h(x) - h(a)\rvert < \varepsilon$	$\lvert x - a\rvert < \varepsilon/2 \to \lvert h(x) - h(a)\rvert < \varepsilon$	$\lvert x - a\rvert < \delta \to \lvert h(x) - h(a)\rvert < \varepsilon$
$a + b = b + a$	$\forall y\,(z \cdot y = y \cdot (a + b))$	$\forall y\,((a + b) \cdot y = y \cdot (a + b))$	$\forall y\,((b + a) \cdot y = y \cdot (a + b))$

In the first row the $f(c)$ inside $R(f(c), c)$ is written out in φ and stays; in the third the hole sits under $\forall y$, and both terms are free for z because neither contains y.

Not an instance. From $2 = 1 + 1$ and the line $2^5 = 32$, the line $(1 + 1)^5 = 3(1 + 1)$, which treats the digit 2 of 32 as an occurrence of the term 2, is not written by =E: the numeral 32 is one constant, with no occurrence of the term 2 inside it; $\varphi = (z^5 = 32)$ gives only $(1 + 1)^5 = 32$.

5.7.7 The rule, phrase and Lean table, completed

Chapter 1 began a table that pairs each rule with the sentence that performs it in a written proof and, for readers who use a proof assistant, with the Lean code that does the same (Sections 1.2–1.6), and Chapter 2 added the rows for negation, the biconditional, and the classical rules (Sections 2.1–2.5). Sections 5.3–5.6 added a row for each quantifier rule, and this section adds the two rows for equality. Here is the whole table for the calculus of this volume.

Rule	Shape	You write
$\wedge$I	φ, $\psi \Longrightarrow \varphi \wedge \psi$	"We show P, and we show Q."
$\wedge$E	$\varphi \wedge \psi \Longrightarrow \varphi$ (or ψ)	"In particular, P."
$\to$E	$\varphi \to \psi$, $\varphi \Longrightarrow \psi$	"Since P implies Q, and P holds, Q."
$\to$I	box φ … $\psi \Longrightarrow \varphi \to \psi$	"Assume P. … Hence Q."
R	φ (available) $\Longrightarrow \varphi$, copied	"As shown above, P."
$\vee$I	$\varphi \Longrightarrow \varphi \vee \psi$ (or $\psi \vee \varphi$)	"So P or Q holds."
$\vee$E	$\varphi \vee \psi$, box φ … χ, box ψ … $\chi \Longrightarrow \chi$	"Case 1: P. … so R. Case 2: Q. … so R. In either case, R."
$\neg$E	φ, $\neg\varphi \Longrightarrow \bot$	"…, which contradicts …"
$\bot$E	$\bot \Longrightarrow$ any φ	"This case cannot occur." "So there is nothing to prove."
$\neg$I	box φ … $\bot \Longrightarrow \neg\varphi$	"Suppose φ. … This is a contradiction. Hence not φ."
$\leftrightarrow$I	box φ … ψ, box ψ … φ $\Longrightarrow \varphi \leftrightarrow \psi$	"($\Rightarrow$) Suppose P. … so Q. ($\Leftarrow$) Suppose Q. … so P."
$\leftrightarrow$E	$\varphi \leftrightarrow \psi$, $\varphi \Longrightarrow \psi$ (or with ψ, φ)	"Since P holds if and only if Q does, and P holds, so does Q."
RAA	box $\neg\varphi$ … $\bot \Longrightarrow \varphi$	"Suppose, for a contradiction, that φ fails. … This is impossible. Hence φ."
$\forall$E	$\forall x\, \varphi \Longrightarrow \varphi[t/x]$	"In particular, for $x = t$, …"; "applying the hypothesis to t, …"
$\forall$I	box $[x_0]$ … $\varphi[x_0/x] \Longrightarrow \forall x\, \varphi$	"Let x be arbitrary. … Since x was arbitrary, …"
$\exists$I	$\varphi[t/x] \Longrightarrow \exists x\, \varphi$	"Take $x = t$. Then …"
$\exists$E	$\exists x\, \varphi$, box $[x_0]$ $\varphi[x_0/x]$ … χ $\Longrightarrow \chi$	"Choose x with … . Then …"

Rule	Shape	You write
=I	nothing $\Longrightarrow t = t$	(usually unwritten) "$t = t$."
=E	$t_1 = t_2$, $\varphi[t_1/x] \Longrightarrow \varphi[t_2/x]$	"Substituting t_2 for t_1, ..."; a chain $t_1 = t_2 = \cdots = t_n$

Remark (For Lean users). ∧I `constructor`; ∧E `h.1`, `h.2`; →E `exact h hp`; →I `intro hp`; R `exact hp`; ∨I `left`, `right` (or `Or.inl hp`); ∨E `Or.elim h` (or `cases h`); ¬E `exact hn hp` (goal `False`); ⊥E `exfalso`; ¬I `intro hp` (goal `False`); ↔I `constructor`; ↔E `h.mp hp`, `h.mpr hq`; RAA `by_contra h`; ∀E `h t` (or `specialize h t`); ∀I `intro x`; ∃I `use t` (or `Exists.intro t h`); ∃E `obtain` (or `Exists.elim h`); =I `rfl`; =E `rw [h]`, chains `calc`. Nothing later depends on these.

Of the derived rules of Chapter 2, MT, CP and DS have their rows in Section 2.3, and ¬¬E and LEM in Section 2.5; the derived sequents that searches use most are collected in Section 3.1. With the two equality rows, every rule of the calculus has its phrase.

Warning (Not allowed: =E). Each step breaks the clause of the reading of Definition 5.7.1, or the warning, shown.

- "From $c = f(b)$ and $Q(f(b))$, write $Q(c)$": the equation is read left to right and rewrites c into $f(b)$; turn it round first, by (sym) (clause 2).
- "From $x = 0$ and $\forall x\,(x \geq x)$, write $\forall x\,(0 \geq 0)$": only free occurrences are replaced, and both x's of $x \geq x$ are bound (the warning "A bound variable is not a place to substitute").
- "From $b = z + 1$ and $\forall z\,(z \leq b)$, write $\forall z\,(z \leq z + 1)$": with $\varphi = \forall z\,(z \leq w)$, the z of $z + 1$ is captured (clause 4).
- "Rewrite a term inside 'it is obvious that ...'": that is not a formula (the warning "Words about knowledge are not formulas").

Remark (Looking ahead). Chapter 8 turns the "You write" column of this table into a full dictionary of proof phrases.

5.7.8 Exercises

5.7.8.1 A. Check your understanding

Exercise. A1

(a) State =I and =E as schemas, with the side condition of =E.

(b) In the derivation below, name for each of lines 4, 5, 6, and 8 the rule, the cited lines, and the formula φ (with a variable z as its hole).

1	$a = b + 1$	
2	$b = 2$	
3	$a \cdot a = a \cdot a$	=I
4	$a \cdot a = (b + 1) \cdot a$	=E 1, 3
5	$a \cdot a = (b + 1) \cdot (b + 1)$	=E 1, 4
6	$a \cdot a = (2 + 1) \cdot (2 + 1)$	=E 2, 5
7	$(2 + 1) \cdot (2 + 1) = 9$	arith
8	$a \cdot a = 9$	=E 7, 6

(c) In the same derivation, could lines 4 and 5 be replaced by one line? With which φ?

(d) True or false, with a reason: from the lines $a = b$ and $P(b)$, one application of =E gives $P(a)$.

(e) True or false, with a reason: from $a = b$ and the line $\forall x\,(x = a \to P(x))$, one application of =E gives $\forall x\,(x = b \to P(x))$.

Solution. *(a)* =I: at any line k, write $t = t$, for any term t, citing nothing. =E: from $t_1 = t_2$ on line i and $\varphi[t_1/x]$ on line j, both available at line k, write $\varphi[t_2/x]$ on line k; side condition: t_1 and t_2 are both free for x in φ.

(b) Line 4: =E 1, 3, with $\varphi = (a \cdot a = z \cdot a)$, which rewrites the first a on the right. Line 5: =E 1, 4, with $\varphi = (a \cdot a = (b + 1) \cdot z)$. Line 6: =E 2, 5, with $\varphi = (a \cdot a = (z + 1) \cdot (z + 1))$, which rewrites both b's at once. Line 8: =E 7, 6, with $\varphi = (a \cdot a = z)$. In each case the left-hand side $a \cdot a$ is written out in φ, so it is kept.

(c) Yes: =E 1, 3 with $\varphi = (a \cdot a = z \cdot z)$ gives $a \cdot a = (b + 1) \cdot (b + 1)$ in one step, rewriting both a's on the right and keeping the two on the left.

(d) False. With the equation $a = b$, =E passes from $\varphi[a/x]$ to $\varphi[b/x]$. The line $P(b)$ is $\varphi[b/x]$ for $\varphi = P(x)$, so the rule would need the equation $b = a$. One application of symmetry first (Proposition 5.7.1 (a)) gives $b = a$, and then =E gives $P(a)$.

(e) True. Take $\varphi = \forall x\,(x = z \to P(x))$: the cited line is $\varphi[a/z]$ and the new line is $\varphi[b/z]$. The terms a and b are constants, so both are free for z, although the hole lies under $\forall x$.

Exercise. A2: Drill

Each part takes a line or two. The parts run from single steps, through a missing line and a justification column, to judging steps and short derivations. Letters a, b, c, d are constants.

(a) What does =I cite?

(b) From $a = b$ and $P(a, a)$, box the holes to get $P(a, b)$.

(c) From $c = d + 1$ and the line $R(c, c \cdot c)$, write the =E line that rewrites only the last c.

(d) From $x_0 = 2$ and $x_0 + x_0 = 4$, which φ gives $2 + x_0 = 4$?

(e) Line 6 is $f(b) + b = 7$, justified "=E 2, 5", and line 2 is $a = b$. Give two possible lines 5.

(f) Justify each line: 1 $a = b$, premise; 2 $\forall x\,(P(x) \to x = c)$, premise; 3 $P(a)$, premise; 4 $P(a) \to a = c$; 5 $a = c$; 6 $a = a$; 7 $b = a$; 8 $b = c$.

(g) From $a = b$ and $P(b)$, does one =E give $P(a)$?

(h) From $a = y$ and $\exists y\,(y < a)$, is $\exists y\,(y < y)$ a legal =E line?

(i) From $x_0 = 0$ and the line $\forall y\,(x_0 \leq y)$, is $\forall y\,(0 \leq y)$ a legal =E line? From $y = 0$ and the same line, is $\forall y\,(x_0 \leq 0)$?

(j) Derive $a = b \vdash f(b) = f(a)$.

(k) Derive $a = b,\ b = c \vdash a = c$ in one step.

(l) Derive $\forall x\,R(x, x) \vdash \forall x\,\forall y\,(x = y \to R(x, y))$.

Solution.

(a) Nothing.

(b) Box the second a only: $\varphi = P(a, z)$.

(c) $R(c, c \cdot (d + 1))$, with $\varphi = R(c, c \cdot z)$.

(d) $\varphi = (z + x_0 = 4)$.

(e) Any two of $f(a) + a = 7$ ($\varphi = (f(z) + z = 7)$), $f(b) + a = 7$ ($\varphi = (f(b) + z = 7)$), $f(a) + b = 7$ ($\varphi = (f(z) + b = 7)$); and the degenerate $f(b) + b = 7$ itself.

(f) 4 $\forall$E 2 (at a); 5 $\to$E 4, 3; 6 =I; 7 =E 1, 6 ($\varphi = (z = a)$); 8 =E 5, 7 ($\varphi = (b = z)$).

(g) No, clause 2: the equation rewrites a into b. Turn it round first (sym), then =E.

(h) No, clause 4: the y put in would be captured by $\exists y$.

(i) The first is legal: $\varphi = \forall y\,(z \leq y)$, and 0 is free for z. The second is not: with $\varphi = \forall y\,(x_0 \leq z)$, the term y of the equation is not free for z in φ (clause 4); the y would be captured.

(j) 2 $f(a) = f(a)$, =I; 3 $f(b) = f(a)$, =E 1, 2, with $\varphi = (f(z) = f(a))$.

(k) 3 $a = c$, =E 2, 1, with $\varphi = (a = z)$.

(l) A flagged box x_0, inside it a flagged box y_0: 2 $x_0 = y_0$, assumption; 3 $R(x_0, x_0)$, $\forall$E 1; 4 $R(x_0, y_0)$, =E 2, 3 ($\varphi = R(x_0, z)$); 5 $x_0 = y_0 \to R(x_0, y_0)$, $\to$I 2–4; 6 $\forall y\,(x_0 = y \to R(x_0, y))$, $\forall$I 2–5; 7 $\forall x\,\forall y\,(x = y \to R(x, y))$, $\forall$I 2–6. (A reflexive relation contains equality.)

5.7.8.2 B. Practice

Exercise. B1: The identity is unique

A set with an operation $\cdot$ has an **identity** e if $e \cdot x = x$ and $x \cdot e = x$ for every x. Derive

$$\forall x\,(e \cdot x = x \wedge x \cdot e = x),\ \ \forall x\,(e' \cdot x = x \wedge x \cdot e' = x) \ \vdash\ e = e',$$

where e and e' are constants. Which half of "there is exactly one identity" does this prove?

Hint: look at the one element that both premises can say something about.

Solution. The goal is an equation between e and e', and the only term that both premises can connect to both is the product $e \cdot e'$. Instantiate the first premise at e' and the second

at e ($\forall$E, reading the terms off the goal):

1	$\forall x\,(e \cdot x = x \wedge x \cdot e = x)$	
2	$\forall x\,(e' \cdot x = x \wedge x \cdot e' = x)$	
3	$e \cdot e' = e' \wedge e' \cdot e = e'$	$\forall$E 1
4	$e \cdot e' = e'$	$\wedge$E 3
5	$e' \cdot e = e \wedge e \cdot e' = e$	$\forall$E 2
6	$e \cdot e' = e$	$\wedge$E 5
7	$e = e'$	=E 6, 4

Line 7 uses the equation $e \cdot e' = e$ of line 6 to rewrite the left side of line 4, with $\varphi = (z = e')$: $\varphi[e \cdot e'/z]$ is line 4, and $\varphi[e/z]$ is $e = e'$. In prose: $e = e \cdot e' = e'$, the first equation because e' is an identity and the second because e is.

This is the "at most one" half: any two identities are equal. It says nothing about whether an identity exists. Section 5.8 shows how such an argument combines with existence.

Exercise. B2: Rewrite, and turn round

A search for $a = b,\ b = c,\ P(a) \vdash P(c) \wedge c = a$ has reached the stage below. *Name the move, then complete it, giving for each* =E *line its holes.*

1	$a = b$	
2	$b = c$	
3	$P(a)$	
	$\vdots$	
9	$P(c) \wedge c = a$	?

Solution. The move is the search step "An equation" of this section: mark the holes, and turn the equation round with an =I line if it points the wrong way.

1	$a = b$	
2	$b = c$	
3	$P(a)$	
4	$P(b)$	$=$E 1, 3
5	$P(c)$	$=$E 2, 4
6	$a = c$	$=$E 2, 1
7	$a = a$	$=$I
8	$c = a$	$=$E 6, 7
9	$P(c) \wedge c = a$	$\wedge$I 5, 8

Lines 4 and 5: $\varphi = P(z)$. Line 6: $\varphi = (a = z)$, so line 1 is $\varphi[b/z]$ and the equation $b = c$ refills the hole with c. Line 8: $\varphi = (z = a)$; line 7 is $\varphi[a/z]$, and the equation $a = c$ refills the box with c. Lines 7–8 are (sym) written out; with the abbreviation of the text, line 8 is "$c = a$ (sym) 6".

Exercise. B3: Derive or break

For each claim, derive it with the rules of Chapters 1–2 and of this chapter, or break it with a counter-reading.

(a) $f(a) = b,\ a = c \vdash f(c) = b$.
(b) $f(a) = f(b) \vdash a = b$.
(c) $\forall x\,(f(x) = x) \vdash f(f(c)) = c$.
(d) $\forall x \forall y\,(f(x) = f(y) \to x = y),\ f(a) = f(b) \vdash a = b$.
(e) $a \neq b,\ b \neq c \vdash a \neq c$.
(f) $a = b,\ b \neq c \vdash a \neq c$.
(g) $f(a) = a \vdash f(f(a)) = a$.
(h) $f(f(a)) = a \vdash f(a) = a$.

Solution.

(a) Derive. One $=$E, with $\varphi = (f(z) = b)$:

1	$f(a) = b$	
2	$a = c$	
3	$f(c) = b$	$=$E 2, 1

(b) Break. Counter-reading: the domain $\{0, 1\}$, f sending both to 0, a naming 0 and b naming 1. Then $f(a) = 0 = f(b)$, and $a = b$ is false.

(c) Derive. Instantiate at the compound term $f(c)$ as well as at c:

1	$\forall x\,(f(x) = x)$	
2	$f(f(c)) = f(c)$	$\forall$E 1
3	$f(c) = c$	$\forall$E 1
4	$f(f(c)) = c$	=E 3, 2

Line 4 has $\varphi = (f(f(c)) = z)$.

(d) Derive.

1	$\forall x\,\forall y\,(f(x) = f(y) \to x = y)$	
2	$f(a) = f(b)$	
3	$\forall y\,(f(a) = f(y) \to a = y)$	$\forall$E 1
4	$f(a) = f(b) \to a = b$	$\forall$E 3
5	$a = b$	$\to$E 4, 2

(e) Break. Counter-reading: the domain $\{0, 1\}$, a and c naming 0, b naming 1. Then $a \neq b$ and $b \neq c$ are true, and $a \neq c$ is false: two different letters, one object.

(f) Derive.

1	$a = b$	
2	$b \neq c$	
3	$\quad a = c$	asm, for $\neg$I
4	$\quad b = c$	=E 1, 3
5	$\quad \bot$	$\neg$E 4, 2
6	$a \neq c$	$\neg$I 3–5

Line 4 has $\varphi = (z = c)$.

(g) Derive.

1	$f(a) = a$	
2	$f(f(a)) = f(f(a))$	=I
3	$f(f(a)) = f(a)$	=E 1, 2
4	$f(f(a)) = a$	=E 1, 3

Line 3 has $\varphi = (f(f(a)) = f(z))$, and line 4 has $\varphi = (f(f(a)) = z)$.

(h) Break. Counter-reading: the domain $\{0, 1\}$, f exchanging 0 and 1, a naming 0. Then $f(f(0)) = f(1) = 0$, so the premise is true, and $f(0) = 1$, so $f(a) = a$ is false.

(i) and (d) differ by the premise that f is one-to-one; (e) and (f) look alike, and $\neq$ is not transitive, since different letters may name one object; (g) and (h) are converses. Worked example types: Proposition 5.7.1; the chain derivation of "Calculations are chains of substitutions"; the congruence derivation.

Marking scheme. (10 points) 1 for each of the eight claims (a derivation with every φ named, or a counter-reading with the domain and every symbol's value); 2 for the closing comparison, with the point about (e) and (f) stated. Fatal errors (0 for the claim): an =E with the equation read right to left; a counter-reading in which two letters that must differ (by a premise) name one object.

5.7.8.3 C. Going deeper

Exercise. C1: Two different objects

(a) Derive $\exists x\, \exists y\, (x \neq y) \vdash \forall x\, \exists y\, (x \neq y)$, and name the line that uses a classical rule.

(b) Read both formulas aloud. Derive the converse, or break it.

Solution.

(a) Read off the goal and the premise. The premise is an existential (twice), so both $\exists$E boxes open first, with names x_0 and y_0. The goal is universal, so a flagged box u_0 opens inside them. The new goal $\exists y\, (u_0 \neq y)$ is postponed, and no available line decides whether u_0 is x_0, so the search splits by LEM:

1		$\exists x\, \exists y\, (x \neq y)$	
2	x_0	$\exists y\, (x_0 \neq y)$	asm, for $\exists$E
3	y_0	$x_0 \neq y_0$	asm, for $\exists$E
4	u_0	$u_0 = x_0 \vee u_0 \neq x_0$	LEM
5		$u_0 = x_0$	asm, for $\vee$E
6		$u_0 = y_0$	asm, for $\neg$I
7		$x_0 = y_0$	=E 5, 6
8		$\bot$	$\neg$E 7, 3
9		$u_0 \neq y_0$	$\neg$I 6–8
10		$\exists y\, (u_0 \neq y)$	$\exists$I 9
11		$u_0 \neq x_0$	asm, for $\vee$E
12		$\exists y\, (u_0 \neq y)$	$\exists$I 11
13		$\exists y\, (u_0 \neq y)$	$\vee$E 4, 5–10, 11–12
14		$\forall x\, \exists y\, (x \neq y)$	$\forall$I 4–13
15		$\forall x\, \exists y\, (x \neq y)$	$\exists$E 2, 3–14
16		$\forall x\, \exists y\, (x \neq y)$	$\exists$E 1, 2–15

Line 7 is =E with $\varphi = (z = y_0)$: line 6 is $\varphi[u_0/z]$, and the equation $u_0 = x_0$ refills the hole with x_0. The classical line is line 4, LEM, a derived rule whose derivation uses RAA (Section 2.5). At line 14, u_0 appears in no premise, in neither open assumption (lines 2 and 3) and not in line 14. At lines 15 and 16 the witnesses y_0 and x_0 appear neither in the last lines, nor in the existentials, nor in the open assumptions.

(b) The premise says "there are two different objects"; the conclusion says "every object has one different from it". The converse derives: 2 $\exists y\, (x \neq y)$, $\forall$E 1 (at x); 3 $\exists x\, \exists y\, (x \neq y)$, $\exists$I 2 (witness x). Over a one-element domain both are false; over larger ones both are true.

Marking scheme. (10 points) 2: two $\exists$E boxes with different names, each aimed at the goal it must reach. 2: a flagged box u_0 for the $\forall$ (inside or outside the $\exists$E boxes), and the goal postponed. 3: the split on $u_0 = x_0$, with the first case reaching $u_0 \neq y_0$ by $\neg$I and =E,

and the second giving the witness x_0 at once. 1: the classical line named, with the reason. 2: (b), the converse derived, with both readings. An RAA box on $\neg\exists y\,(u_0 \neq y)$ in place of the LEM split earns full marks if correct, its RAA line then named as the classical one. Fatal errors (0 for the step): the flag u_0 chosen as x_0; an $\exists$E box closed on a line that mentions its witness.

5.8 Unique Existence, Counting and Descriptions

Mathematics says "the limit", "the maximum", "the inverse", and it defines new words all the time: "we define $|x|$ to be ...", "call n even if ...". This section is about the logic behind both habits. It starts with "there is exactly one", which licenses "the". Then it counts with equality, turns "the x such that φ" into a name that may be used only after uniqueness is proved, and sets a function symbol beside its graph. Last, it shows why introducing a new symbol by a definition never proves anything new about the old ones. No new rule is needed: equality and the quantifiers do all the work.

5.8.1 Unique existence

Mathematics says "**the** limit of the sequence", "**the** supremum of A", "**the** additive inverse of a". Each "the" makes two claims. There is such an object, and there is only one, so that the name picks it out. The second claim is the one that is easy to forget, and the one the claim at the start of Section 5.7 was about.

"There is exactly one x with φ": there is one, and any other is it.

Definition 5.8.1. Unique Existence

Let φ be a formula, x a variable, and y a variable **other than x that does not appear in φ** (Section 5.2). Then

$$\exists!x\,\varphi \quad \text{abbreviates} \quad \exists x\,(\varphi \wedge \forall y\,(\varphi[y/x] \to y = x)).$$

We read $\exists!x\,\varphi$ as "there is **exactly one** x such that φ", or "there is a **unique** x with φ".

In words: there is an x that satisfies φ, and every y that satisfies φ (that is, satisfies φ with y in place of x) is equal to that x. The first conjunct is existence. The second is uniqueness, and it is a statement about the x already found: any other candidate coincides with it. Like the bounded quantifiers of Definition 5.1.1, $\exists!$ is not a new symbol of the language but shorthand, and every rule applies to the formula it abbreviates.

Well-definedness. The abbreviation does not depend on the choice of y, up to renaming a bound variable. The argument, in the fold below, rests on one fact about substitution that the proof of Proposition 5.8.1 uses again: *if y does not appear in φ, then $(\varphi[y/x])[t/y]$ is $\varphi[t/x]$.*

↻ **Second reading** · *Well-definedness, proved*

This is a structural induction on φ of the kind in Section 5.2. In a term, the y's of $s[y/x]$ are

exactly the former x's, so $(s[y/x])[t/y] = s[t/x]$; atomic formulas, $\bot$ and the connectives follow clause by clause. For $\forall x\,\psi$, both sides are $\forall x\,\psi$, since substitution for x stops at $\forall x$ and y does not appear in ψ (the second fact of Section 5.2). For $\forall z\,\psi$ with z not x, also z is not y, because y does not appear in φ; so both substitutions pass under the quantifier and the induction hypothesis applies, and $\exists$ is the same. Now let y and y' be two admissible choices. If they are the same variable there is nothing to show. Otherwise y' does not appear in $\varphi[y/x] \to y = x$, since it is neither x nor y and does not appear in φ; so the bound y of $\forall y\,(\varphi[y/x] \to y = x)$ may be renamed y', which gives $\forall y'\,((\varphi[y/x])[y'/y] \to y' = x)$. By the fact with $t = y'$, this is $\forall y'\,(\varphi[y'/x] \to y' = x)$, the formula written with y'.

Examples. Simplest first; each is read over the domain named, with the informal truth of Section 5.1.

1. Over $\mathbb{R}$, $\exists!x\,(x + 2 = 5)$ is true. *Existence:* $3 + 2 = 5$. *Uniqueness:* if $y + 2 = 5$, then subtracting 2 gives $y = 3$, so y is the x already found.
2. Over $\mathbb{N}$, $\exists!x\,(x \cdot x = 4)$ is true. *Existence:* $2 \cdot 2 = 4$. *Uniqueness:* for a natural number y, if $y \leq 1$ then $y \cdot y \leq 1$, and if $y \geq 3$ then $y \cdot y \geq 9$; so $y \cdot y = 4$ forces $y = 2$.
3. Over $\mathbb{R}$, $\exists!x\,\forall y\,(x + y = y)$ is true: 0 is the **unique** additive identity. *Existence:* $0 + y = y$ for every y. *Uniqueness:* suppose $z + y = y$ for every y. Apply this to the one instance $y = 0$ (by $\forall$E): $z + 0 = 0$. Since also $z + 0 = z$, symmetry and transitivity give $z = 0$. The whole uniqueness proof is the choice of the instance $y = 0$.

Nonexample by minimal change. Keep example 2 and change the domain to $\mathbb{Z}$. Existence still holds, with the same witness 2. Uniqueness fails: $y = -2$ also satisfies $y \cdot y = 4$, and $-2 \neq 2$. So over $\mathbb{Z}$, $\exists!x\,(x \cdot x = 4)$ is false, and the failing clause is exactly the second conjunct.

Quick check. Is $\exists!x\,(x < 1)$ true over $\mathbb{N}$? Over $\mathbb{Z}$?

Solution. Over $\mathbb{N}$, yes: $0 < 1$, and every natural number below 1 is 0. Over $\mathbb{Z}$, no: existence holds, and uniqueness fails, since 0 and -1 are both below 1.

↻ **Second reading** · *Why y must not appear, and a degenerate case*

Why y must not appear in φ. The condition keeps the new quantifier $\forall y$ from capturing anything of φ's own. Take $\varphi = (x \leq y)$, in which y is free, and read over $\mathbb{N}$ with y naming 0. "There is exactly one natural number x with $x \leq 0$" is true: $x = 0$. The formula written with the forbidden bound variable y is

$$\exists x\,\big(x \leq y \wedge \forall y\,(y \leq y \to y = x)\big),$$

because $(x \leq y)[y/x]$ is $y \leq y$. Its second conjunct says that **every** natural number equals x, which is false. The free y of φ, which was to name 0, has been captured by $\forall y$. With an admissible variable, say z, the formula is $\exists x\,\big(x \leq y \wedge \forall z\,(z \leq y \to z = x)\big)$, which is true.

Degenerate. For any constant c, $\exists!x\,(x = c)$ holds in every domain, and it is derivable from no premises (the derivation follows). By contrast, $\exists!x\,(x = x)$ says that the domain has exactly one element, which is false in $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{R}$: existence holds and uniqueness fails.

The derivation of $\exists!x\,(x = c)$, with the abbreviation written out, shows that $\exists!$ needs no rules of its own:

1	$c = c$	=I
2	y_0 | $y_0 = c$	asm, for $\to$I
3	$y_0 = c \to y_0 = c$	$\to$I 2–2
4	$\forall y\,(y = c \to y = c)$	$\forall$I 2–3
5	$c = c \wedge \forall y\,(y = c \to y = c)$	$\wedge$I 1, 4
6	$\exists x\,(x = c \wedge \forall y\,(y = c \to y = x))$	$\exists$I 5

Line 3 closes a box whose only line is its assumption: the box shows $y_0 = c \to y_0 = c$, as $P \to P$ is shown in Chapter 1. Line 6 is $\exists$I with the witness c, and it reads line 5 as $\psi[c/x]$ for $\psi = (x = c \wedge \forall y\,(y = c \to y = x))$, which marks with x the first c and the last one, and keeps the other two. The choice of which occurrences to mark is the one discussed with Definition 5.5.1.

Warning. **Uniqueness is not existence.** A uniqueness argument ("suppose y also satisfies φ; then $y = x$") proves nothing about whether anything satisfies φ. Over $\mathbb{R}$, every two solutions of $x^2 = -1$ are equal, vacuously, because there are none; yet no real number is "the" square root of -1. Conversely, an existence proof does not give uniqueness: $x^2 = 1$ has the solution 1, and "the solution" would be wrong, since -1 is another. Before writing "the", check both halves.

5.8.2 How uniqueness is used and proved

A proof of "exactly one" rarely follows the shape of Definition 5.8.1, which ties the uniqueness clause to the witness just found. It proves two separate statements: there is at least one, and there is at most one. The second is "any two objects with the property are equal", and it makes no mention of a witness. The next result says that the two shapes say the same thing.

Proposition 5.8.1. Two Halves of Unique Existence

Let φ be a formula, x a variable, and y a variable other than x that does not appear in φ. Then

$$\exists!x\,\varphi \;\dashv\vdash\; \exists x\,\varphi \wedge \forall x\,\forall y\,(\varphi \wedge \varphi[y/x] \to x = y).$$

The right-hand side is "existence and at most one". Its second conjunct is what a mathematician proves under the heading "uniqueness": take two objects x and y with the property, and show $x = y$.

Idea. Both directions are searches in the style of Section 3.1 with the quantifier steps. For $\vdash$, the premise is an $\exists$, so open its $\exists$E box first, with a witness x_0: inside it, $\varphi[x_0/x]$ gives existence at once by $\exists$I, and "at most one" is a $\forall\forall$ goal, which opens two flagged boxes u_0, v_0. Two objects u_0, v_0 with the property are each equal to the witness x_0, by the uniqueness clause, and so equal to each other: the witness is the "third point" that links them. For $\dashv$,

open the ∃E box on $\exists x\,\varphi$ with a witness x_0; the goal inside is $\exists! x\,\varphi$, whose witness will be x_0 again, and the uniqueness clause for x_0 comes from "at most one" applied to the pair x_0, y_0.

The two derivations that carry this out are in the fold below.

Second reading · *The proof of the two halves*

Proof. Choose four distinct variables x_0, u_0, v_0, y_0 that do not appear in φ, and none of which is x or y. Three facts about substitution are used throughout. First, since y does not appear in φ, $(\varphi[y/x])[t/y]$ is $\varphi[t/x]$ for every term t, the fact proved for the well-definedness of Definition 5.8.1. Second, $\varphi[y/x]$ has no free x (every free x became y), so it is unchanged by a substitution for x. Third, for a term t that does not contain y, $\varphi[t/x]$ contains no y, so it is unchanged by a substitution for y.

($\vdash$) The following is a derivation with premise $\exists! x\,\varphi$, written out.

1	$\exists x\,(\varphi \wedge \forall y\,(\varphi[y/x] \to y = x))$	
2	x_0 | $\varphi[x_0/x] \wedge \forall y\,(\varphi[y/x] \to y = x_0)$	asm, for ∃E
3	$\varphi[x_0/x]$	∧E 2
4	$\forall y\,(\varphi[y/x] \to y = x_0)$	∧E 2
5	$\exists x\,\varphi$	∃I 3
6	u_0 | v_0 | $\varphi[u_0/x] \wedge \varphi[v_0/x]$	asm, for →I
7	$\varphi[u_0/x]$	∧E 6
8	$\varphi[v_0/x]$	∧E 6
9	$\varphi[u_0/x] \to u_0 = x_0$	∀E 4
10	$u_0 = x_0$	→E 9, 7
11	$\varphi[v_0/x] \to v_0 = x_0$	∀E 4
12	$v_0 = x_0$	→E 11, 8
13	$v_0 = v_0$	=I
14	$x_0 = v_0$	=E 12, 13
15	$u_0 = v_0$	=E 14, 10
16	$\varphi[u_0/x] \wedge \varphi[v_0/x] \to u_0 = v_0$	→I 6–15
17	$\forall y\,(\varphi[u_0/x] \wedge \varphi[y/x] \to u_0 = y)$	∀I 6–16
18	$\forall x\,\forall y\,(\varphi \wedge \varphi[y/x] \to x = y)$	∀I 6–17
19	$\exists x\,\varphi \wedge \forall x\,\forall y\,(\varphi \wedge \varphi[y/x] \to x = y)$	∧I 5, 18
20	$\exists x\,\varphi \wedge \forall x\,\forall y\,(\varphi \wedge \varphi[y/x] \to x = y)$	∃E 1, 2–19

Lines 9 and 11 are ∀E on line 4 at u_0 and at v_0, by the first substitution fact; each term is a variable not bound in $\varphi[y/x] \to y = x_0$, so it is free for y there. Lines 13 and 14 are the symmetry of Proposition 5.7.1 (a), written out; line 15 is =E with the equation $x_0 = v_0$ and $\psi = (u_0 = z)$, which turns line 10 into $u_0 = v_0$. Line 17 is ∀I with the flag v_0: line 16 is its instance at v_0 (by the third and first substitution facts), and v_0 appears in no premise, in no assumption still open at line 17 (only line 2 is, and v_0 is not in it), and not in line 17. Line 18 is ∀I with the flag u_0 in the same way; its instance at u_0 is line 17, by the second substitution fact. Finally, line 20 closes the ∃E box: its last line 19 does not contain x_0, and x_0 is in neither line 1 nor any open assumption.

(⊣) The following is a derivation with the right-hand side as premise.

1	$\exists x\, \varphi \land \forall x\, \forall y\, (\varphi \land \varphi[y/x] \to x = y)$	
2	$\exists x\, \varphi$	∧E 1
3	$\forall x\, \forall y\, (\varphi \land \varphi[y/x] \to x = y)$	∧E 1
4	x_0 | $\varphi[x_0/x]$	asm, for ∃E
5	y_0 | $\varphi[y_0/x]$	asm, for →I
6	$\forall y\, (\varphi[x_0/x] \land \varphi[y/x] \to x_0 = y)$	∀E 3
7	$\varphi[x_0/x] \land \varphi[y_0/x] \to x_0 = y_0$	∀E 6
8	$\varphi[x_0/x] \land \varphi[y_0/x]$	∧I 4, 5
9	$x_0 = y_0$	→E 7, 8
10	$x_0 = x_0$	=I
11	$y_0 = x_0$	=E 9, 10
12	$\varphi[y_0/x] \to y_0 = x_0$	→I 5–11
13	$\forall y\, (\varphi[y/x] \to y = x_0)$	∀I 5–12
14	$\varphi[x_0/x] \land \forall y\, (\varphi[y/x] \to y = x_0)$	∧I 4, 13
15	$\exists x\, (\varphi \land \forall y\, (\varphi[y/x] \to y = x))$	∃I 14
16	$\exists x\, (\varphi \land \forall y\, (\varphi[y/x] \to y = x))$	∃E 2, 4–15

Line 6 is ∀E at x_0 on line 3, by the second substitution fact, and line 7 is ∀E at y_0 on line 6, by the third and first. Lines 10 and 11 are symmetry, written out. Line 13 is ∀I with the flag y_0: line 12 is its instance at y_0, by the first substitution fact; at line 13 the open assumptions are line 4, which does not contain y_0, the premise does not contain it either, and neither does line 13. Line 15 is ∃I with the witness x_0, which is free for x since it does not appear in φ; line 14 is its instance at x_0, by the second substitution fact. Line 16 closes the ∃E box: line 15 does not contain x_0, and x_0 is in neither line 2 nor the premise, which is the only assumption open outside the box. This proves both directions.

■

In practice this is how "there is exactly one" is proved: an existence proof, then a separate "suppose x and y both satisfy φ; then $x = y$".

One more instance: "the maximum". A **maximum** of a set A of reals is an element $m \in A$ with $a \leq m$ for every $a \in A$. If m and m' are both maxima, then $m' \leq m$ (apply the property of m to the element $m' \in A$) and $m \leq m'$ (apply the property of m' to $m \in A$), so $m = m'$ by the order fact that $a \leq b$ and $b \leq a$ give $a = b$. That is "at most one"; a set may have no maximum at all, such as $(0, 1)$, so existence is a separate question. Only when both halves are known may we write "the maximum of A".

5.8.3 Counting with equality

"At least one" is $\exists$, and "at most one" is the second half of Proposition 5.8.1. Beyond one, counting needs the sign $\neq$.

Definition 5.8.2. Counting with Equality

Let φ be a formula, x a variable, and y, z two different variables, other than x, that do not appear in φ.

- "**At least two** x satisfy φ" abbreviates $\exists x\, \exists y\, (x \neq y \wedge \varphi \wedge \varphi[y/x])$.
- "**At most one** x satisfies φ" abbreviates $\forall x\, \forall y\, (\varphi \wedge \varphi[y/x] \to x = y)$.
- "**At most two** x satisfy φ" abbreviates $\forall x\, \forall y\, \forall z\, (\varphi \wedge \varphi[y/x] \wedge \varphi[z/x] \to x = y \vee x = z \vee y = z)$.
- "**Exactly two** x satisfy φ" abbreviates the conjunction of "at least two" and "at most two".

For a number $n \geq 2$, "at least n" uses n different variables $x_1, \ldots, x_n$ that do not appear in φ, and asks for all $n(n-1)/2$ inequalities $x_i \neq x_j$ together with φ at each; "at most n" says that among any $n+1$ objects satisfying φ, two are equal; "exactly n" is both.

In words: two letters count two objects only when the formula says that they differ, and "at most two" says that three letters satisfying φ cannot name three different objects. As with $\exists!$, these are abbreviations, not new symbols, and $\exists!$ stays the only counting shorthand.

Example. Exactly Two Square Roots of One

Over $\mathbb{R}$, show that exactly two x satisfy $x \cdot x = 1$.

Solution. At least two: the witnesses 1 and -1: $1 \neq -1$, $1 \cdot 1 = 1$ and $(-1) \cdot (-1) = 1$.

At most two: let u, v, w be real numbers with $u \cdot u = v \cdot v = w \cdot w = 1$. Each is 1 or -1, since $y \cdot y = 1$ gives $(y-1)(y+1) = 0$. If $u = v$, the disjunction $u = v \vee u = w \vee v = w$ holds. Otherwise one of u, v is 1 and the other -1, and w, being 1 or -1, equals one of them. Three letters, at most two objects: that is what "at most two" means.

Warning. **Different letters, perhaps the same object.** $\exists x\, \exists y\, (x \cdot x = 1 \wedge y \cdot y = 1)$ is true over $\mathbb{R}$ with x and y both naming 1: it says no more than $\exists x\, (x \cdot x = 1)$. Two letters count two objects only when the formula says $x \neq y$. The same holds in derivations: two $\exists$E boxes with the names x_0 and y_0 may have named the same object, and nothing gives $x_0 \neq y_0$.

Quick check. Write "at least two primes are even" with Prime and Even, and decide it over $\mathbb{N}$.

Solution. $\exists x\,\exists y\,(x \neq y \wedge (\mathrm{Prime}(x) \wedge \mathrm{Even}(x)) \wedge (\mathrm{Prime}(y) \wedge \mathrm{Even}(y)))$. It is false: an even prime is divisible by 2, so it is 2, and any two even primes are equal.

Only, *except* and *the smallest* are counting words too; the drill at the end of the section translates them.

5.8.4 The x such that φ

Once there is exactly one x with φ, mathematics names it: "the maximum of A", "the additive inverse of a", "$\sqrt{2}$". The name is then used like a constant, because the proof has earned it.

Definition 5.8.3. Definite Description

Let Γ be a list of formulas and φ a formula whose only free variable is x, and suppose $\Gamma \vdash \exists!x\,\varphi$. **Introducing a name for the x such that φ** means taking a constant c that appears neither in Γ nor in φ, and letting derivations from Γ use the **defining line** $\varphi[c/x]$ as one more premise. The phrase "the x such that φ" names c.

In words, and why both halves are needed: existence is what makes the defining line safe, and the proposition at the end of this section shows that it proves nothing new about the old symbols. Uniqueness is what makes the word "the" honest. If c and c' are both introduced for φ, then $\Gamma, \varphi[c/x], \varphi[c'/x] \vdash c = c'$: take the "at most one" half of Proposition 5.8.1, instantiate it at c and at c' ($\forall$E twice, with the substitution facts used in the proof of Proposition 5.8.1), and fire it by $\wedge$I and $\to$E. So every proof that names the x such that φ names the same object.

Russell's unpacking. A sentence about the x such that φ can be unpacked into one that names nothing. Write $\mathrm{Max}(A, m)$ for "m is a maximum of A", that is, $m \in A \wedge \forall a \in A\,(a \leq m)$ (the paragraph "One more instance" above). "The maximum of A is less than 2" unpacks to

$$\exists m\,(\mathrm{Max}(A, m) \wedge \forall y\,(\mathrm{Max}(A, y) \to y = m) \wedge m < 2)\,:$$

there is exactly one maximum, and it is less than 2.

Two readings of a negation. When nothing satisfies φ, the unpacked sentence is false, and a negation can be read two ways. The interval $A = (0, 1)$ has no maximum: an $m \in A$ is exceeded by $(m + 1)/2$, which is also in A. Read "the maximum of A is not 1" with the negation inside, $\exists m\,(\mathrm{Max}(A, m) \wedge \forall y\,(\mathrm{Max}(A, y) \to y = m) \wedge m \neq 1)$: it is false, since there is no m. Read it with the negation outside, $\neg\exists m\,(\mathrm{Max}(A, m) \wedge \forall y\,(\mathrm{Max}(A, y) \to y = m) \wedge m = 1)$: it is true. The two readings agree whenever there is exactly one maximum, which is why "the" waits for $\exists!$. The same trap sits in "$\lim a_n \neq 0$" for a sequence with no limit, and analysis writes instead "the sequence converges, and its limit is not 0".

Quick check. Over $\mathbb{R}$, may a proof speak of "the x with $x \cdot x = 2$"? Of "the $x \geq 0$ with $x \cdot x = 2$"?

Solution. The first, no: $\sqrt{2}$ and $-\sqrt{2}$ both have square 2, so uniqueness fails. The second, yes: existence is the school fact of Section 5.1 ($\sqrt{2}$ exists), and uniqueness holds, since two non-negative numbers with the same square are equal (if $0 \leq u < v$, then $u \cdot u < v \cdot v$).

5.8.5 A function symbol and its graph

A function symbol gives every object exactly one value, since every term names an object (Section 5.2). A two-place predicate $G(x, y)$ may or may not behave like that. Call G the **graph of a function** when $\forall x\, \exists! y\, G(x, y)$: every x has exactly one partner. Then a new function symbol f may be introduced with the defining line $\forall x\, G(x, f(x))$, and $f(t)$ is "the y such that $G(t, y)$": the name of the last subsection, with x as a parameter.

Over $\mathbb{R}$, $G(x, y)$: $y = x \cdot x + 1$ is the graph of a function (existence: $y = x \cdot x + 1$ itself; uniqueness: two numbers equal to $x \cdot x + 1$ are equal). $G(x, y)$: $y \cdot y = x$ is not: $x = -1$ has no partner, and $x = 4$ has two, 2 and -2. For $x \geq 0$, the predicate $y \geq 0 \wedge y \cdot y = x$ gives exactly one partner, and the function it defines is $\sqrt{x}$, written only under the hypothesis $x \geq 0$ (the fold on partial terms in Section 5.2).

So "a function" and "a relation in which every object has exactly one partner" are two ways of saying the same thing; the function symbol is the shorter.

Quick check. Over $\mathbb{Z}$, is $x = 2 \cdot y$ the graph of a function, with y as the value at x? Is $y = 2 \cdot x$?

Solution. The first is not: $x = 1$ has no partner, since no integer y has $2y = 1$. The second is: each x has exactly one partner, $2x$; the function is $f(x) = 2x$.

5.8.6 Definitions add words, not facts

"An integer n is **even** if $n = 2k$ for some integer k" adds a predicate symbol with its **definition** $\forall x\, (\mathrm{Even}(x) \leftrightarrow \exists k\, (x = 2 \cdot k))$. "We define $|x|$ to be x if $x \geq 0$ and $-x$ otherwise" adds a function symbol. Its graph $G(x, y)$: $(x \geq 0 \wedge y = x) \vee (x < 0 \wedge y = -x)$ gives each x exactly one partner (one of $x \geq 0$, $x < 0$ holds and gives its y, and the two cases exclude each other), and $|\cdot|$ is introduced by $\forall x\, G(x, |x|)$. A definition is a new premise, and premises prove things. The next result says what a definition can prove: nothing new about the old symbols.

Proposition 5.8.2. Definitions Prove Nothing New

Let Γ be a list of formulas and ψ a formula.

(a) Let P be a one-place predicate symbol that appears neither in Γ nor in ψ, and φ a formula in which P does not appear and whose only free variable is x. If $\Gamma, \forall x\, (P(x) \leftrightarrow \varphi) \vdash \psi$, then $\Gamma \vdash \psi$.

(b) Let φ be a formula whose only free variable is x, and c a constant that appears in none of Γ, φ, ψ. If $\Gamma \vdash \exists x\, \varphi$ and $\Gamma, \varphi[c/x] \vdash \psi$, then $\Gamma \vdash \psi$.

In words: whatever is proved with a defined symbol, about symbols that were there before it, could have been proved without it. A definition makes proofs shorter, never stronger.

Part (b) asks only for existence; uniqueness, as the subsection on "the x such that φ" said, is what makes "the" name one object. The same holds for predicates with several places, and for a function symbol f introduced by $\forall x\, G(x, f(x))$ once $\forall x\, \exists! y\, G(x, y)$ is proved; we do not prove those cases, and nothing later depends on them.

Idea. For (b), rename c everywhere into a variable that appears nowhere, and put the whole derivation inside an $\exists$E box on $\exists x\, \varphi$. For (a), replace every atomic formula $P(t)$ by $\varphi[t/x]$; the definition becomes $\forall x\, (\varphi \leftrightarrow \varphi)$, which needs no premise. The fold below carries both out.

↻ **Second reading** · *Why definitions prove nothing new*

Proof.

(b) Let D be a derivation of ψ from Γ and $\varphi[c/x]$. Choose a variable x_0 that appears nowhere in D, Γ, φ or ψ (a derivation from Γ need not write every member of Γ), and replace c by x_0 everywhere in D. Every line stays a correct line: the rules treat constants and variables alike as terms; x_0 is not a flag or a witness name of D, since it appears nowhere there; and it is bound nowhere, so every "free for" condition still holds. Γ and ψ do not contain c and do not change, and the premise $\varphi[c/x]$ becomes $\varphi[x_0/x]$, since c is not in φ. Now take a derivation of $\exists x\, \varphi$ from Γ, open an $\exists$E box (Definition 5.6.1) with the name x_0 and the assumption $\varphi[x_0/x]$, copy the changed D inside it, and close the box on ψ: by the choice of x_0, it is not in ψ, not in $\exists x\, \varphi$, and not in Γ, which holds the only premises.

(c) Let D be a derivation of ψ from Γ and $\forall x\, (P(x) \leftrightarrow \varphi)$. Rename the bound variables of φ to letters that appear nowhere in D (the convention of Section 5.2), and replace every atomic formula $P(t)$ in D by $\varphi[t/x]$. Because x is the only free variable of φ and its bound letters are new, the replacement commutes with every substitution a rule makes and changes no free variable of any line; so every line stays a correct line, and every side condition still holds. Γ and ψ do not contain P and do not change; the definition becomes $\forall x\, (\varphi \leftrightarrow \varphi)$, which is derived from no premises by a flagged box, with a flag x_0 that appears nowhere in Γ, φ or D, ending in $\varphi[x_0/x] \leftrightarrow \varphi[x_0/x]$, got by $\leftrightarrow$I from two one-line boxes. Writing that derivation in front of the changed D derives ψ from Γ. (After the renaming the defining premise is $\forall x\, (\varphi' \leftrightarrow \varphi)$, with φ' the renamed φ, which the convention of Section 5.2 treats as $\forall x\, (\varphi \leftrightarrow \varphi)$.)

■

What a definition sounds like. "We define ..." and "call n even if ..." introduce a symbol with its defining line; the right-hand side must not mention the new symbol, and a function needs exactly one value for each argument. "Let L be the limit of (a_n)" introduces a name by a description, and may be written only after it has been proved that there is exactly one such L. Every "the" carries that debt; "a limit" does not.

Remark (Looking ahead). Chapter 8 collects "we define ..." and "the ..." with the other phrases of written proof.

5.8.7 Exercises

5.8.7.1 A. Check your understanding

Exercise. A1

(a) Write $\exists! x\,(x \cdot x = 9)$ without the symbol $\exists!$, and decide whether it is true over $\mathbb{N}$ and over $\mathbb{Z}$.

(b) Write "exactly two integers x satisfy $x \cdot x = 9$" by Definition 5.8.2, and decide it over $\mathbb{Z}$.

(c) When may a proof write "the x such that φ", and what may it then use about that object?

(d) State in words the two parts of Proposition 5.8.2.

(e) Exercise B1 of Section 5.7 showed that any two identities of an operation are equal. Explain how to turn its derivation into the second conjunct of Proposition 5.8.1 for $\varphi = \forall z\,(x \cdot z = z \wedge z \cdot x = z)$.

Solution. *(a)* $\exists x\,(x \cdot x = 9 \wedge \forall y\,(y \cdot y = 9 \to y = x))$. Over $\mathbb{N}$ it is true: $x = 3$ works, and a natural number y with $y \cdot y = 9$ is 3, since $y \leq 2$ gives $y \cdot y \leq 4$ and $y \geq 4$ gives $y \cdot y \geq 16$. Over $\mathbb{Z}$ it is false: whatever x is proposed, 3 and -3 both satisfy $y \cdot y = 9$, and they cannot both equal x.

(b) $\exists x\,\exists y\,(x \neq y \wedge x \cdot x = 9 \wedge y \cdot y = 9) \wedge \forall x\,\forall y\,\forall z\,(x \cdot x = 9 \wedge y \cdot y = 9 \wedge z \cdot z = 9 \to x = y \vee x = z \vee y = z)$. True: $3 \neq -3$ are two solutions, and every solution is 3 or -3 (since $(y-3)(y+3) = 0$), so among three solutions two are equal.

(c) After $\exists! x\,\varphi$ has been proved from the premises in hand. It may use the defining line $\varphi[c/x]$, and only what follows from it with the premises.

(d) Part (a): adding a predicate symbol by a definition $\forall x\,(P(x) \leftrightarrow \varphi)$ proves nothing new that does not mention P. Part (b): naming a witness of a proved existential by a new constant proves nothing new that does not mention the constant.

(e) Take $\varphi = \forall z\,(x \cdot z = z \wedge z \cdot x = z)$, and run lines 3–7 again with variables u_0, v_0 in place of e, e': inside flagged boxes for u_0 and v_0 and a $\to$I box on $\varphi[u_0/x] \wedge \varphi[v_0/x]$, the two premises (with z as their bound variable) come from that assumption by $\wedge$E, and the same steps give $u_0 = v_0$. Then $\to$I and $\forall$I twice give $\forall x\,\forall y\,(\varphi \wedge \varphi[y/x] \to x = y)$, the second conjunct of the proposition. With the existence of one identity as well, Proposition 5.8.1 gives "there is exactly one".

Exercise. A2: Drill

Each part takes a line or two. The parts run from writing counting statements, through judging them and the counting words, to descriptions and graphs.

(a) Write $\exists! x\,(x+1=3)$ without $\exists!$.

(b) Write "at least two elements of A are negative".

(c) Write "at most two x satisfy $P(x)$".

(d) Write "exactly two x satisfy $P(x)$" as one formula of the shape $\exists x\,\exists y\,(\cdots \wedge \forall z\,(\dots))$.

(e) Does $\exists x\,\exists y\,(P(x)\wedge P(y))$ say "at least two x satisfy $P(x)$"?

(f) Over $\mathbb{N}$, translate "0 is the only natural number n with $n+n=n$", and decide it.

(g) Translate "every prime except 2 is odd". Does your formula say that 2 is not odd?

(h) Write "m is the smallest element of A", and say what must be proved before writing "the smallest element of A".

(i) In $A=\{2,3,4\}$, ordered by divisibility, which elements are maximal (no other element of A is a multiple of them), and is there a maximum (an element that is a multiple of every element of A)?

(j) $A=(0,1]$ has no smallest element. Read "the smallest element of A is not positive" both ways, and decide each.

(k) Over $\mathbb{R}$, does $x\cdot y=1$ give each x exactly one partner y? Over the non-zero reals?

(l) A proof says "let c be the real number with $c+3=5$". What must have been proved, and what may be used about c?

Solution.

(a) $\exists x\,(x+1=3\wedge\forall y\,(y+1=3\to y=x))$.

(b) $\exists x\,\exists y\,(x\neq y\wedge(x\in A\wedge x<0)\wedge(y\in A\wedge y<0))$.

(c) $\forall x\,\forall y\,\forall z\,(P(x)\wedge P(y)\wedge P(z)\to x=y\vee x=z\vee y=z)$.

(d) $\exists x\,\exists y\,(x\neq y\wedge P(x)\wedge P(y)\wedge\forall z\,(P(z)\to z=x\vee z=y))$.

(e) No: x and y may name the same object; it says only $\exists x\,P(x)$.

(f) $0+0=0\wedge\forall x\,(x+x=x\to x=0)$. True: $x+x=x$ gives $x=0$ by cancelling x.

(g) $\forall x\,(\mathrm{Prime}(x)\wedge x\neq 2\to\mathrm{Odd}(x))$. It says nothing about 2. If "except" is meant to say that 2 is not odd as well, add $\wedge\neg\mathrm{Odd}(2)$.

(h) $m\in A\wedge\forall a\in A\,(m\leq a)$. Existence only: two smallest elements are each $\leq$ the other, hence equal, so uniqueness is automatic.

(i) 3 and 4 are maximal; 2 is not, since 4 is a multiple of 2. There is no maximum: 3 is not a multiple of 2 or of 4, and 4 is not a multiple of 3; and 2 is not a multiple of 3 either. Maximal and maximum differ.

(j) With $\mathrm{Min}(A,m)$ the formula of (h): negation inside, $\exists m\,(\mathrm{Min}(A,m)\wedge\forall y\,(\mathrm{Min}(A,y)\to y=m)\wedge\neg(m>0))$, false, since there is no m. Negation outside, $\neg\exists m\,(\mathrm{Min}(A,m)\wedge\forall y\,(\mathrm{Min}(A,y)\to y=m)\wedge m>0)$, true. (For $m\in A$, $m/2\in A$ is smaller.)

(k) No: $x=0$ has none, since $0\cdot y=0$. Over $x\neq 0$, yes: the partner is $1/x$, and $x\cdot y=x\cdot y'=1$ gives $y=y'$.

(l) $\exists! x\,(x+3=5)$: existence makes the name safe, uniqueness makes "the" right. It may use the defining line $c+3=5$, and what follows from it, such as $c=2$.

5.8.7.2 B. Practice

Exercise. B1: Exactly two, by halves

Over $\mathbb{N}$, show that exactly two natural numbers x satisfy $x \cdot x < 3$: write the statement by Definition 5.8.2, then prove "at least two" and "at most two" separately.

Solution. The statement is

$$\exists x\, \exists y\, (x \neq y \wedge x \cdot x < 3 \wedge y \cdot y < 3) \wedge \forall x\, \forall y\, \forall z\, (x \cdot x < 3 \wedge y \cdot y < 3 \wedge z \cdot z < 3 \to x = y \vee x = z \vee y = z).$$

At least two: 0 and 1: $0 \neq 1$, $0 \cdot 0 = 0 < 3$ and $1 \cdot 1 = 1 < 3$.

At most two: if $x \cdot x < 3$ then $x \leq 1$, since $x \geq 2$ gives $x \cdot x \geq 4$; so $x \in \{0, 1\}$. Three natural numbers u, v, w with squares below 3 take values in $\{0, 1\}$, so two of them are equal. Worked example type: Example (Exactly Two Square Roots of One).

Marking scheme. (10 points) 3: the formula, with $x \neq y$ and the three-way disjunction. 3: "at least two", with both witnesses and their difference. 4: "at most two", with $x \leq 1$ justified and the pigeonhole step stated. Fatal error (0 for the half): "at least two" written without $x \neq y$.

Exercise. B2: A unique square root of zero

Over $\mathbb{R}$, write $\exists! x\, (x \cdot x = 0)$ without $\exists!$, and prove it: existence, then uniqueness in the form "suppose y also satisfies ...".

Hint: a product of two non-zero reals is non-zero.

Solution. Written out, the statement is $\exists x\, (x \cdot x = 0 \wedge \forall y\, (y \cdot y = 0 \to y = x))$. We take the witness $x = 0$ and prove the two conjuncts.

Existence. $0 \cdot 0 = 0$.

Uniqueness. Let y be a real number, and suppose $y \cdot y = 0$. If $y \neq 0$, then $y \cdot y$ is a product of two non-zero reals and so is non-zero, which contradicts $y \cdot y = 0$. Hence $y = 0$, which is the witness. Since y was arbitrary, every y with $y \cdot y = 0$ equals 0.

So 0 witnesses the existential, and $\exists! x\, (x \cdot x = 0)$ holds. The rules behind the words: "we take the witness $x = 0$" is $\exists$I, postponed to the end as usual; "let y be a real number" is a flagged box for $\forall$I; "suppose $y \cdot y = 0$" is a $\to$I box; "if $y \neq 0$ ... contradicts" is an RAA box on $y \neq 0$. (That box gives $y = 0$ directly; $\neg$I would give only $\neg(y \neq 0)$. Since for reals $y = 0$ or $y \neq 0$ is also a school fact, a proof by cases would do as well.)

Exercise. B3: Exactly one fixed point

A **fixed point** of a function $f : \mathbb{R} \to \mathbb{R}$ is a real x with $f(x) = x$.

(a) Formalize "f has exactly one fixed point" over $\mathbb{R}$, first with $\exists!$ and then without it.

(b) Write a formula without $\exists!$ that says "f does **not** have exactly one fixed point", in the

form "there is none, or there are two different ones".

(c) Decide the statement of (a) for $f(x) = x \cdot x$, for $f(x) = x + 1$ and for $f(x) = 2 \cdot x$, and check your answer against the formula of (b).

Solution. *(a)* With $\exists!$: $\exists! x\,(f(x) = x)$. Written out, with y as the new variable: $\exists x\,(f(x) = x \wedge \forall y\,(f(y) = y \to y = x))$.

(b) $\neg\exists x\,(f(x) = x) \vee \exists x\,\exists y\,(f(x) = x \wedge f(y) = y \wedge x \neq y)$. Read informally over $\mathbb{R}$: either no real number is fixed, or two different ones are. That is exactly the failure of "exactly one". If no x is fixed, the first disjunct holds. If some x is fixed but the uniqueness clause fails for it, then some y is fixed and differs from x. Conversely, either disjunct rules out a unique fixed point: with none there is no witness, and two different fixed points $x \neq y$ would both have to equal any proposed witness w, giving $x = y$. (Section 6.2 derives such negations mechanically, one quantifier at a time.)

(c) For $f(x) = x \cdot x$: $x \cdot x = x$ holds for $x = 0$ and for $x = 1$, so there are two different fixed points; (a) is false, and the second disjunct of (b) is true with $x = 0$, $y = 1$. For $f(x) = x + 1$: $x + 1 = x$ has no solution, so (a) is false and the first disjunct of (b) is true. For $f(x) = 2 \cdot x$: $2x = x$ gives $x = 0$, and 0 is fixed; so (a) is true with the witness 0, and both disjuncts of (b) are false.

Exercise. B4: The empty set is unique

Read the domain as consisting of sets, and let E be the axiom of extensionality, $\forall x\,\forall y\,(\forall z\,(z \in x \leftrightarrow z \in y) \to x = y)$.

(a) Derive $E,\ \forall z\,\neg(z \in a),\ \forall z\,\neg(z \in b) \vdash a = b$.

(b) Which half of "there is exactly one set with no elements" does (a) prove, and what would give the other half?

(c) Suppose also $\exists x\,\forall z\,\neg(z \in x)$. Which result of this section gives $\exists! x\,\forall z\,\neg(z \in x)$, and what may a proof then write?

Solution.

(a) The goal $a = b$ is the consequent of an instance of E; its hypothesis $\forall z\,(z \in a \leftrightarrow z \in b)$ is built in a flagged box, each direction of the biconditional by $\perp$E.

1	$\forall x\,\forall y\,(\forall z\,(z \in x \leftrightarrow z \in y) \to x = y)$	
2	$\forall z\,\neg(z \in a)$	
3	$\forall z\,\neg(z \in b)$	
4	z_0 | $z_0 \in a$	asm, for $\leftrightarrow$I
5	$\neg(z_0 \in a)$	$\forall$E 2
6	$\bot$	$\neg$E 4, 5
7	$z_0 \in b$	$\bot$E 6
8	$z_0 \in b$	asm, for $\leftrightarrow$I
9	$\neg(z_0 \in b)$	$\forall$E 3
10	$\bot$	$\neg$E 8, 9
11	$z_0 \in a$	$\bot$E 10
12	$z_0 \in a \leftrightarrow z_0 \in b$	$\leftrightarrow$I 4–7, 8–11
13	$\forall z\,(z \in a \leftrightarrow z \in b)$	$\forall$I 4–12
14	$\forall y\,(\forall z\,(z \in a \leftrightarrow z \in y) \to a = y)$	$\forall$E 1
15	$\forall z\,(z \in a \leftrightarrow z \in b) \to a = b$	$\forall$E 14
16	$a = b$	$\to$E 15, 13

(b) "At most one": any two sets without elements are equal. The other half, "at least one", needs a premise of its own, an axiom saying that a set without elements exists: from E alone the claim $\exists x\,\forall z\,\neg(z \in x)$ breaks. Counter-reading: the domain $\{0\}$, with $0 \in 0$. E is true, since there is only one object and so $x = y$ holds for every x, y; and $\exists x\,\forall z\,\neg(z \in x)$ is false, since the only object, 0, has the element 0.

(c) Run (a) with flags x_0, y_0 in place of a, b, inside a $\to$I box on $\forall z\,\neg(z \in x_0) \wedge \forall z\,\neg(z \in y_0)$ ($\wedge$E gives the two premises), and close with $\to$I and $\forall$I twice: that is the second conjunct of Proposition 5.8.1 for $\varphi = \forall z\,\neg(z \in x)$. With the existence premise, the direction $\dashv$ of Proposition 5.8.1 gives $\exists! x\,\forall z\,\neg(z \in x)$, and by Definition 5.8.3 a proof may now say "**the** empty set" and name it $\emptyset$. Worked example types: Example (Equal Sets Have the Same Elements) (Section 5.7); the two halves of Proposition 5.8.1.

Marking scheme. (10 points) 4: (a), the biconditional built by $\leftrightarrow$I from two boxes, each closed by $\bot$E, then $\forall$I, two $\forall$E on E and $\to$E. 3: (b), "at most one" (1), existence as a separate premise (1), and the counter-reading with E checked true and the existence claim false (1). 3: (c), the flags and the $\to$I box, the right direction of Proposition 5.8.1, and the license to write "the". Fatal error (0 for (a)): $a = b$ "by extensionality" with no instance of E written.

Exercise. B5: Composites of one-to-one maps

A function symbol f is **one-to-one** if $\forall x\,\forall y\,(f(x) = f(y) \to x = y)$. Derive

$$\forall x\,\forall y\,(f(x) = f(y) \to x = y),\ \ \forall x\,\forall y\,(g(x) = g(y) \to x = y) \ \vdash\ \forall x\,\forall y\,(f(g(x)) = f(g(y)) \to x = y),$$

then write the proof in two sentences.

Solution.

1	$\forall x\,\forall y\,(f(x) = f(y) \to x = y)$	
2	$\forall x\,\forall y\,(g(x) = g(y) \to x = y)$	
3	x_0 | y_0 | $f(g(x_0)) = f(g(y_0))$	asm, for $\to$I
4	$\forall y\,(f(g(x_0)) = f(y) \to g(x_0) = y)$	$\forall$E 1
5	$f(g(x_0)) = f(g(y_0)) \to g(x_0) = g(y_0)$	$\forall$E 4
6	$g(x_0) = g(y_0)$	$\to$E 5, 3
7	$\forall y\,(g(x_0) = g(y) \to x_0 = y)$	$\forall$E 2
8	$g(x_0) = g(y_0) \to x_0 = y_0$	$\forall$E 7
9	$x_0 = y_0$	$\to$E 8, 6
10	$f(g(x_0)) = f(g(y_0)) \to x_0 = y_0$	$\to$I 3–9
11	$\forall y\,(f(g(x_0)) = f(g(y)) \to x_0 = y)$	$\forall$I 3–10
12	$\forall x\,\forall y\,(f(g(x)) = f(g(y)) \to x = y)$	$\forall$I 3–11

Lines 4 and 5 instantiate the premise for f at the compound terms $g(x_0)$ and $g(y_0)$: the move *instantiate at the term you need*. In prose: "Let x and y satisfy $f(g(x)) = f(g(y))$. Since f is one-to-one, $g(x) = g(y)$, and since g is one-to-one, $x = y$."

Marking scheme. (10 points) 2: the two flags and the $\to$I box. 4: the $\forall$E steps at $g(x_0)$, $g(y_0)$, then at x_0, y_0, each with $\to$E. 2: the closing $\forall$I lines with the flag checks. 2: the prose, one sentence per use of one-to-one. Fatal error (0 for the derivation): the premise for f instantiated at x_0, y_0.

5.8.7.3 C. Going deeper

Exercise. C1: Two unique quantifiers do not commute

Show that $\exists!x\,\exists!y\,R(x,y)$ and $\exists!y\,\exists!x\,R(x,y)$ can have different truth values, by giving a finite domain D and a two-place relation R on it for which the first is true and the second false. Check both, reading $\exists!$ informally.

Hint: first say in words what each sentence claims about the rows and columns of a table of R.

Solution. **What any answer must satisfy.** Call x *single* if exactly one y has $R(x,y)$, and call y *co-single* if exactly one x has $R(x,y)$. The first sentence says that **exactly one** element is single. The second says that exactly one element is co-single. So an answer needs exactly one single element and a number of co-single elements other than one: none, or two or more.

An example. Take $D = \{0,1,2\}$ and $R = \{(1,2),(2,0),(2,1)\}$.

- Single elements: 0 has no partner, 1 has exactly one (2), and 2 has two (0 and 1). So 1 is the only single element, and $\exists!x\,\exists!y\,R(x,y)$ is true, with the witness 1.
- Co-single elements: $y=0$ is reached only from $x=2$; $y=1$ only from $x=2$; $y=2$ only from $x=1$. All three are co-single, so there is not exactly one, and $\exists!y\,\exists!x\,R(x,y)$ is false.

No example exists on one or two elements (there are only $2+16=18$ relations on such domains). On one element, 0 is single and co-single exactly when $R(0,0)$ holds, so the two sentences agree. On two elements, suppose exactly one element is single; after relabeling, call it 0, and let c be its only partner. Then 1 is not single, so it has no partners or two. If it has none, column c has one entry and the other column has none. If it has two, column c has two entries and the other column has one. Either way exactly one element is co-single. Exchanging the roles of rows and columns gives the converse, so the two sentences have the same truth value. Three elements is therefore the smallest possible domain.

The moral: $\exists!$ is not a quantifier that commutes with its own kind, as two $\exists$'s do. "Exactly one" hides a $\forall$ (the uniqueness clause), and the order of a $\forall$ and an $\exists$ matters.

Exercise. C2: "At most one", two ways

Let P be a one-place predicate symbol. Derive both directions of

$$\forall x\,\forall y\,(P(x)\wedge P(y)\to x=y)\ \dashv\vdash\ \neg\exists x\,\exists y\,((P(x)\wedge P(y))\wedge x\neq y),$$

and say which of your two derivations uses RAA, and at which line.

Solution (Hint 1). The right-hand side is a negation: for $\vdash$, open a $\neg$I box, and inside it two $\exists$E boxes. For $\dashv$, the goal is $\forall\forall$, then an implication, then an equation.

Solution (Hint 2). In the $\dashv$ direction, open flagged boxes for x_0 and y_0 and a $\to$I box on $P(x_0)\wedge P(y_0)$. The goal is now the atomic formula $x_0=y_0$. Which resources can give it?

Solution. ($\vdash$) The goal is a negation, so a $\neg$I box opens on $\exists x\,\exists y\,(\dots)$; inside, the premise-directed step opens an $\exists$E box for each $\exists$, with the goal $\bot$ as each box's last line. The premise, instantiated at the two witnesses, gives $x_0=y_0$, which clashes with $x_0\neq y_0$.

1	$\forall x\,\forall y\,(P(x) \land P(y) \to x = y)$		
2	$\quad \exists x\,\exists y\,((P(x) \land P(y)) \land x \neq y)$		asm, for ¬I
3	$\quad x_0 \mid \exists y\,((P(x_0) \land P(y)) \land x_0 \neq y)$		asm, for ∃E
4	$\quad\quad y_0 \mid (P(x_0) \land P(y_0)) \land x_0 \neq y_0$		asm, for ∃E
5	$\quad\quad\quad P(x_0) \land P(y_0)$		∧E 4
6	$\quad\quad\quad x_0 \neq y_0$		∧E 4
7	$\quad\quad\quad \forall y\,(P(x_0) \land P(y) \to x_0 = y)$		∀E 1
8	$\quad\quad\quad P(x_0) \land P(y_0) \to x_0 = y_0$		∀E 7
9	$\quad\quad\quad x_0 = y_0$		→E 8, 5
10	$\quad\quad\quad \bot$		¬E 9, 6
11	$\quad\quad \bot$		∃E 3, 4–10
12	$\quad \bot$		∃E 2, 3–11
13	$\neg\exists x\,\exists y\,((P(x) \land P(y)) \land x \neq y)$		¬I 2–12

The brackets in the statement are written out because the default groups a longer conjunction to the right (Section 4.5); with the grouping $(P(x_0) \land P(y_0)) \land x_0 \neq y_0$ on line 4, ∧E takes out its two halves on lines 5 and 6. At line 11, y_0 is not in $\bot$, not in line 3, not in the premise, and not in the assumptions still open (lines 2 and 3). At line 12 the same holds for x_0, with line 2 the only open assumption.

(¬) The goal-directed steps open flagged boxes for x_0 and y_0 and a →I box on $P(x_0) \land P(y_0)$, leaving the goal $x_0 = y_0$. No resource gives an equation, so the search is stuck, and step 5 of the method applies: an RAA box on $x_0 \neq y_0$. Inside, ∧I and two ∃I steps rebuild the formula that the premise denies.

1	$\neg\exists x\,\exists y\,((P(x) \land P(y)) \land x \neq y)$			
2	$x_0 \mid y_0 \mid\quad P(x_0) \land P(y_0)$			asm, for →I
3	$\quad\quad\quad\quad x_0 \neq y_0$			asm, for RAA
4	$\quad\quad\quad\quad (P(x_0) \land P(y_0)) \land x_0 \neq y_0$			∧I 2, 3
5	$\quad\quad\quad\quad \exists y\,((P(x_0) \land P(y)) \land x_0 \neq y)$			∃I 4
6	$\quad\quad\quad\quad \exists x\,\exists y\,((P(x) \land P(y)) \land x \neq y)$			∃I 5
7	$\quad\quad\quad\quad \bot$			¬E 6, 1
8	$\quad\quad\quad x_0 = y_0$			RAA 3–7
9	$\quad\quad P(x_0) \land P(y_0) \to x_0 = y_0$			→I 2–8
10	$\quad \forall y\,(P(x_0) \land P(y) \to x_0 = y)$			∀I 2–9
11	$\forall x\,\forall y\,(P(x) \land P(y) \to x = y)$			∀I 2–10

The ⊣ derivation uses RAA**, at line 8**: the box opens with $x_0 \neq y_0$, that is $\neg(x_0 = y_0)$, and closes with $x_0 = y_0$. Closing it by ¬I instead would give $\neg\neg(x_0 = y_0)$, and no rule of

Chapter 1 or Sections 2.1–2.4 removes the double negation. (With the derived rule ¬¬E of Section 2.5, that ¬I box followed by ¬¬E is an equally good derivation; its classical step is the ¬¬E line, which hides an RAA.) The ⊢ derivation uses no classical rule.

Marking scheme. (10 points) - 2: ⊢: opens a ¬I box on the existential, with the goal ⊥. - 2: ⊢: two ∃E boxes with **fresh** witnesses, each ending in ⊥, and the premise instantiated at those witnesses. - 2: ⊣: two flagged boxes and a →I box, in the order read off the goal. - 2: ⊣: an RAA box on $x_0 \neq y_0$, with the clash obtained by ∧I and two ∃I steps. - 2: names the ⊣ direction as the one using RAA, and the line. - A ¬I box on $x_0 \neq y_0$ followed by ¬¬E, or an LEM split on $x_0 = y_0$ whose second case reaches ⊥ the same way (then ⊥E), earns the 2 points of the RAA-box line; for the naming line, the reader names the ¬¬E or LEM line as the classical step, a derived rule whose derivation uses RAA (Section 2.5). Fatal errors (0 for the step): an ∃E box closing on a line that contains its witness; a ∀I over a name in an open assumption; ¬I claimed to give $x_0 = y_0$.

Exercise. C3: One and one make two

Let P and Q be one-place predicate symbols. Show that from $\exists!x\,P(x)$, $\exists!x\,Q(x)$ and $\neg\exists x\,(P(x) \wedge Q(x))$ it follows that exactly two x satisfy $P(x) \vee Q(x)$, in the form

$$\exists x\,\exists y\,\big(x \neq y \wedge (P(x) \vee Q(x)) \wedge (P(y) \vee Q(y)) \wedge \forall z\,(P(z) \vee Q(z) \to z = x \vee z = y)\big).$$

Give the proof in numbered steps, naming the rule behind each; a full derivation is not required. Which premise does the conjunct $x \neq y$ need, and what happens without it?

Solution.

1. By Proposition 5.8.1, the first premise gives $\exists x\,P(x)$ and "at most one P", and the second gives $\exists x\,Q(x)$ and "at most one Q" (∧E).
2. Open two ∃E boxes early, with different names: $P(a_0)$ and $Q(b_0)$, both aimed at the goal.
3. $a_0 \neq b_0$: in a ¬I box on $a_0 = b_0$, =E turns $P(a_0)$ into $P(b_0)$; ∧I with $Q(b_0)$ and ∃I give $\exists x\,(P(x) \wedge Q(x))$, which ¬E clashes with the third premise.
4. $P(a_0) \vee Q(a_0)$ and $P(b_0) \vee Q(b_0)$ by ∨I.
5. Let z_0 be arbitrary (a flagged box) with $P(z_0) \vee Q(z_0)$ (a →I box), and split by ∨E. If $P(z_0)$: "at most one P" at z_0 and a_0 (∀E twice, ∧I, →E) gives $z_0 = a_0$, and ∨I gives $z_0 = a_0 \vee z_0 = b_0$. If $Q(z_0)$: "at most one Q" at z_0 and b_0 gives $z_0 = b_0$, and ∨I again.
6. →I and ∀I give the last conjunct; ∧I joins the four; ∃I with the witness b_0 for y, then a_0 for x. Both ∃E boxes close, since the goal mentions neither name.

The third premise is used only in step 3. Without it the conclusion breaks: the domain $\{0\}$ with P and Q both true of 0 makes the first two premises true, and only one object satisfies $P(x) \vee Q(x)$.

Marking scheme. (10 points) 2: the two halves of each premise by Proposition 5.8.1. 1: two ∃E boxes with different names. 2: $a_0 \neq b_0$ by ¬I, =E and the third premise. 3: the last conjunct by the flagged box, the →I box and ∨E, each case using the right "at most one". 1: the closing ∃I order and the box conditions. 1: the counter-reading without the third premise. Fatal errors (0 for the step): $a_0 \neq b_0$ asserted because the names differ (the warning

"Different letters, perhaps the same object"); an ∃E box closed on a line that mentions a_0 or b_0.

5.8.7.4 R. Review

Exercise. R1

Consider the terms built from variables, the constant 0 and the two-place function symbol $+$. Show that in every such term the number of occurrences of $+$ is one less than the number of occurrences of variables and constants, counted together. Check it on $(x + 0) + (y + x)$.

Solution. For a term t, let $a(t)$ count the occurrences of variables and constants in t, and $p(t)$ those of $+$. We show $p(t) = a(t) - 1$ for every t, by induction on the clauses of Definition 5.2.1, the method of Theorem 4.3.1 (Section 4.3).

Base: t is a variable or 0. Then $p(t) = 0$ and $a(t) = 1$.

Step: t is $t_1 + t_2$, and the claim holds for t_1 and t_2. Then

$$p(t) = p(t_1) + p(t_2) + 1 = (a(t_1) - 1) + (a(t_2) - 1) + 1 = a(t_1) + a(t_2) - 1 = a(t) - 1.$$

The term $(x + 0) + (y + x)$ has four occurrences of variables and constants and three of $+$. This is the counting argument of Lemma 4.3.2 (Section 4.3), on new data. About 10 minutes.

Marking scheme. (10 points) 2: a and p defined. 3: the base case, both kinds of atomic term. 4: the step, with the hypothesis used for both t_1 and t_2 and the $+1$ for the new symbol. 1: the check. Fatal error (0 for the step): induction on the length of the term with no case for $t_1 + t_2$.

Exercise. R2

Derive $\forall x\,(\neg Q(x) \to \neg P(x)) \vdash \forall x\,(P(x) \to Q(x))$, and name the line that uses a classical rule.

Solution.

1	$\forall x\,(\neg Q(x) \to \neg P(x))$		
2	x_0 $P(x_0)$		asm, for $\to$I
3	$\neg Q(x_0) \to \neg P(x_0)$		$\forall$E 1
4	$\neg Q(x_0)$		asm, for RAA
5	$\neg P(x_0)$		$\to$E 3, 4
6	$\bot$		$\neg$E 2, 5
7	$Q(x_0)$		RAA 4–6
8	$P(x_0) \to Q(x_0)$		$\to$I 2–7
9	$\forall x\,(P(x) \to Q(x))$		$\forall$I 2–8

Line 9 cites the flagged box from line 2, the convention for a flagged box that immediately contains another; x_0 is in no premise, in no open assumption, and not in line 9. Lines 2 and 4–8 are the derivation of Proposition 2.5.2 (Section 2.5) at $P(x_0)$, $Q(x_0)$, with line 3 as its premise. The classical line is line 7, RAA (Definition 2.5.1). About 10 minutes.

Marking scheme. (10 points) 2: the flagged box opened first and closed by $\forall$I with the flag checked. 2: $\forall$E at the flag. 4: the $\to$I box and, inside it, the RAA box on $\neg Q(x_0)$. 2: line 7 named as the classical line. A $\neg$I box on $\neg Q(x_0)$ followed by $\neg\neg$E, or Proposition 2.5.2 used as a derived rule on line 3, earns the 4 points of the box line; its classical line (the $\neg\neg$E line, or the derived-rule line, whose derivation uses RAA) earns the naming 2. Fatal error (0): $\forall$E outside the flagged box at a constant, then $\forall$I over that constant.

Exercise. R3

You have derived $(P \wedge Q) \to R \vdash P \to (Q \to R)$. Derive $\forall x\,((P(x) \wedge Q(x)) \to R(x)) \vdash \forall x\,(P(x) \to (Q(x) \to R(x)))$, and say which lines are the propositional derivation and what the quantifier adds.

Solution.

1	$\forall x\,((P(x) \wedge Q(x)) \to R(x))$		
2	x_0 $(P(x_0) \wedge Q(x_0)) \to R(x_0)$		$\forall$E 1
3	$P(x_0)$		asm, for $\to$I
4	$Q(x_0)$		asm, for $\to$I
5	$P(x_0) \wedge Q(x_0)$		$\wedge$I 3, 4
6	$R(x_0)$		$\to$E 2, 5
7	$Q(x_0) \to R(x_0)$		$\to$I 4–6
8	$P(x_0) \to (Q(x_0) \to R(x_0))$		$\to$I 3–7
9	$\forall x\,(P(x) \to (Q(x) \to R(x)))$		$\forall$I 2–8

Lines 3–8 are the first derivation of Proposition 1.4.2 (Section 1.4) at the statements $P(x_0)$, $Q(x_0)$, $R(x_0)$, with line 2 as its premise. The quantifier adds line 2 (∀E at the flag) and line 9 (∀I), whose side condition holds because the boxes of lines 3 and 4 have closed. About 10 minutes.

Marking scheme. (10 points) 3: the flagged box with ∀E at the flag. 4: the two →I boxes in the order of Proposition 1.4.2. 3: the propositional lines named and the side condition of line 9 checked. Fatal error (0): ∀I written while the box of line 3 or 4 is open.

Exercise. R4

Derive $\forall x\,\forall y\,(f(x) = y \to g(y) = x) \vdash \forall x\,(g(f(x)) = x)$.

Solution.

1		$\forall x\,\forall y\,(f(x) = y \to g(y) = x)$	
2	x_0	$\forall y\,(f(x_0) = y \to g(y) = x_0)$	∀E 1
3		$f(x_0) = f(x_0) \to g(f(x_0)) = x_0$	∀E 2
4		$f(x_0) = f(x_0)$	=I
5		$g(f(x_0)) = x_0$	→E 3, 4
6		$\forall x\,(g(f(x)) = x)$	∀I 2–5

The tools: instantiate at the term you need (Definition 5.3.1, Section 5.3), here the compound term $f(x_0)$ at line 3, and =I (Definition 5.7.1, Section 5.7), which makes the hypothesis of line 3 at no cost. Why the look-alikes fail: =E has no equation to cite, since no line is an equation until =I writes one; ∀E at a fresh variable y_0 leaves the hypothesis $f(x_0) = y_0$, which nothing proves; and there is no existential, so ∃E has nothing to open. About 15 minutes.

Marking scheme. (10 points) 2: the flagged box. 4: the second ∀E, at $f(x_0)$. 3: =I and →E. 1: the closing ∀I with the flag checked. Fatal error (0): ∀E at a variable y_0 followed by a step that treats $f(x_0) = y_0$ as proved.

Chapter 6

Working with Quantifiers: Translation, Laws and Search

A first course in analysis defines "$a_n \to L$" with three quantifiers and two bounds, then asks for a proof that a sequence does **not** converge, which starts from its negation. Chapter 5 gave the rules one per section; this chapter puts them to work: turning a definition into a formula, negating and rearranging it by laws, and finding the derivation a claim needs. *You know every quantifier rule. Can you turn a definition from analysis into a formula, negate it without error, and find the derivation it needs?* It ends where Chapter 5 began: why "uniformly continuous implies continuous" has a short derivation and a false converse.

Remark (Looking back). Chapter 3 did for the connectives what this chapter does for the quantifiers. Section 3.3 turned arguments in words into derivations; Section 6.1 does the same for definitions with quantifiers. Section 3.2 sorted the tasks, and Section 3.1 gave the search method; Section 6.3 does both again, with flags and witnesses. Section 6.2 collects the equivalences the tasks keep using. Section 6.4 has no twin: it is the chapter's capstone.

What you need. Chapters 1–5. From Chapter 3: the search method (Section 3.1) and the four kinds of task (Section 3.2). From Chapter 4: truth tables and semantic entailment (Definition 4.5.3), completeness (Theorem 4.8.4), the move *test the subgoal* (Section 4.7) and the TAUT step for propositional formulas (Definition 4.8.1). From Chapter 5: the four quantifier rules with their side conditions (Definition 5.4.1, Definition 5.6.1), bounded quantifiers and the convention for mixing kinds of number (Definition 5.1.1), substitution, free-for and renaming (Section 5.2), equality (Section 5.7), and counter-readings (Section 5.3). School arithmetic and the order of $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ are used in examples as facts.

Time. About 20 hours on the core track (all of A, one B and one C per section, the exit checks); about 35 hours for everything.

Minimum before *Sets, Relations, and Cardinality*. Chapters 1–6 and 8, core; Sections 7.1–7.2 before its Chapter 3.

How to read this chapter. Every section is on the first reading, with its folds closed. The folds hold the proofs of the bounded rules, of the replacement rule and of TAUT on the skeleton, the derivations of the distribution and passage laws, and short remarks on edge cases. They are not optional: later proofs may use them.

6.0.1 Roadmap

- **Translating mathematics.** Definitions from analysis as formulas, worked from the outside in; the bounded rules; five traps, each shown wrong by a reading; "infinitely many" against "eventually".
- **Negating and moving quantifiers.** "Not every" is "some not"; negation pushed inward by the replacement rule; the four forms A, E, I, O and their square; the laws that move a quantifier across a connective, and prenex form.
- **Finding first-order derivations.** The search method with flags and witnesses; TAUT on the skeleton; the propositional core; the four kinds of task; a library of sequents and a battery.
- **Uniformly continuous implies continuous, rule by rule.** The capstone derivation, the lines where the converse breaks, x^2 as the counter-reading, and witness functions $\delta(\varepsilon)$ against $\delta(x, \varepsilon)$.

6.0.2 Named moves

- **Read it back.** Read a translation back into words before using it. (Section 6.1)
- **Negate quantifier by quantifier.** Push $\neg$ inward, swapping $\forall$ and $\exists$, keeping the bounds. (Section 6.2)
- **Name the witness early, then instantiate at the term you need.** Chapter 5's two moves, in that order. (Section 6.3)
- **The propositional core.** Instantiate at the names in play, reason propositionally, with TAUT, then close the boxes. (Section 6.3)
- **Test the subgoal by a counter-reading.** Section 4.7's move, with quantifiers. (Section 6.3)
- **Read the dependency off the boxes.** A witness may depend only on the flags whose boxes enclose it. (Section 6.4)

Where this goes. *Analysis on the Real Line*, Chapter 1 proves from the real-number axioms the facts about $\mathbb{N}$ in $\mathbb{R}$ and absolute value that Sections 6.1 and 6.2 use as school facts, and it uses "not every" gives "some not" (Section 6.2) in its proof of the Archimedean property. Its Chapters 5 and 6 define continuity and uniform continuity, which Section 6.4 uses only as quantifier patterns. Picking a witness for every ε at once, as one function (Section 6.4), in general appeals to the axiom of choice, which *Sets, Relations, and Cardinality*, Chapter 3 states.

6.0.3 Summary

The bounded rules add no power (Proposition 6.1.1). "Not every" is "some not", with one direction classical (Proposition 6.2.1); a negated implication is a conjunction (Proposition 6.2.2); a bound stays with its quantifier (Proposition 6.2.3). A part may be replaced anywhere by a provably equivalent part, by a syntactic induction (Proposition 6.2.4); a step right for propositional reasons alone is derivable (Proposition 6.3.1). The chapter translated and negated the definitions of analysis, and took the search method and the four tasks to quantifiers. The capstone derived "uniformly continuous implies continuous" (Example (Uniformly continuous implies continuous)) and located the side conditions that stop the converse.

What is open: every "break" here was a counter-reading. Does one rule out every derivation? Chapter 7 asks.

6.0.4 Terms

- **A, E, I, O forms**: every, no, some, some not (Section 6.2).
- **alternation**: a $\forall/\exists$ change in a prefix (Section 6.2).
- **bounded rules**: abbreviations, no new power (Proposition 6.1.1).
- **contradictory, contrary**: exactly one true; never both true (Section 6.2).
- **negated implication**: $\neg(\varphi \to \psi)$ is $\varphi \wedge \neg\psi$ (Proposition 6.2.2).
- **passage laws**: move a quantifier across a connective (Proposition 6.2.7).
- **prenex form**: every quantifier at the front (Section 6.2).
- **propositional core**: instantiate, reason propositionally, close (Section 6.3).
- **propositional skeleton**: prime parts read as letters (Definition 6.3.1).
- **quantifier distribution**: $\forall$ over $\wedge$, $\exists$ over $\vee$ (Proposition 6.2.6).
- **quantifier negation**: "not every" is "some not" (Proposition 6.2.1).
- **replacement of provable equivalents**: swap a part for an equivalent (Proposition 6.2.4).
- **TAUT step, on the skeleton**: propositionally right (Proposition 6.3.1).
- **uniformly continuous**: one δ for all points (Section 6.4).
- **witness function**: a witness depending on flags (Section 6.4).

6.1 Translating Mathematics: Nested and Bounded Quantifiers

The calculus is now complete: Sections 5.3–5.7 gave a rule for using and a rule for proving each quantifier, and the two rules for equality. Before any of those rules can act on a definition from analysis, the definition has to be written as a formula, and written correctly. The definitions of analysis have three or four quantifiers stacked in a fixed order, most of them carrying a bound such as $\varepsilon > 0$ or $n \geq N$, and every bound costs extra lines in a derivation. This section does three things. It translates the standard definitions into formulas, by a procedure that works from the outside in. It turns each quantifier rule into a **bounded** rule that handles the bound in the same step. And it collects the traps that make a translation wrong, each with a reading that shows the wrong formula wrong.

Remark (Looking back). Section 3.3 turned an argument in words into a derivation, with one key for a whole paragraph. This section does the same for definitions with quantifiers. What is new is the order of the quantifiers and the bounds they carry, which the propositional key never had.

6.1.1 Four pieces in one sentence

"The sequence (a_n) converges to L" is written $a_n \to L$, and a first course defines it in one breath: for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that $|a_n - L| < \varepsilon$ for all $n \geq N$. The sentence has three quantifiers, each with a bound ($\varepsilon > 0$, $N \in \mathbb{N}$, $n \in \mathbb{N}$), and a fourth condition, $n \geq N$, hidden in "for all $n \geq N$", which will turn out to sit under an implication. Each piece has a job, and a slip in any piece changes the meaning:

- drop the bound $\varepsilon > 0$, and the challenger may choose $\varepsilon = -1$; then $|a_n - L| < -1$ is impossible, and **no** sequence converges;
- swap "for every ε" and "there is N", and one N must serve every ε; then only sequences

that are eventually equal to L converge;
- read "for all $n \geq N$" as "for all n", and the first terms must already be close to L; then $1, \frac{1}{2}, \frac{1}{3}, \ldots$ converges to nothing.

So translation deserves a procedure, and the rules deserve a form that handles bounds without fuss. We take them in that order.

6.1.2 Translation from the outside in

Find the outermost claim, name its quantifier and its bound, then translate what it claims about each object, in the same way.

The procedure, in three steps:

1. **Fix the domain.** State what the variables range over. The logic of this book has a single domain for all the variables of a formula. Mathematics often mixes kinds of objects in one sentence: in "for every $\varepsilon > 0$ there is an $N \in \mathbb{N}$ such that for every $n \geq N$, ...", the letter ε is a real number, while N and n are natural numbers. The convention is: **the domain is the larger kind** (here $\mathbb{R}$), and **the smaller kind enters as a bound** ($\exists N \in \mathbb{N}$, $\forall n \in \mathbb{N}$). So the sentence begins

$$\forall \varepsilon > 0 \; \exists N \in \mathbb{N} \; \forall n \in \mathbb{N} \; (n \geq N \to \ldots),$$

   and n ranges over real numbers that happen to be required, by the bound, to be natural numbers. A sequence (a_n) becomes a function symbol a, with $a(n)$, written a_n, the n-th term. Its values at real numbers that are not natural numbers are never used, because a is only ever applied to variables bounded by $n \in \mathbb{N}$. Section 5.1 stated this convention in one sentence ("Mixing kinds of number"); here it is in full, and every translation in this chapter follows it.
2. **Find the outermost claim.** Ask what the sentence says first, before any detail: "every $\varepsilon > 0$ has a property", "there is an M with a property", "no prime is largest". That gives the outermost quantifier (or connective) and its bound.
3. **Recurse.** Translate the property, which is a sentence about the object just quantified, by the same two steps. Stop at atomic formulas: equations, inequalities, membership, and predicate symbols such as $\text{Prime}(m)$.

A bound is attached to its own quantifier with the connective of Definition 5.1.1: $\to$ under $\forall$, $\wedge$ under $\exists$. The bounded notation $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ does this automatically, which is why we use it.

A mixed stage. Halfway through, a translation is part words and part symbols, and writing that stage down catches most slips. Take "for every positive ε, some element of A lies within ε of s". Each line below translates one more quantifier and leaves the rest in words, in square brackets, until no words are left:

$$\begin{aligned}
&\forall \varepsilon > 0 \; [\text{some element of } A \text{ lies within } \varepsilon \text{ of } s] \\
&\forall \varepsilon > 0 \; \exists x \in A \; [x \text{ lies within } \varepsilon \text{ of } s] \\
&\forall \varepsilon > 0 \; \exists x \in A \; (|x - s| < \varepsilon)
\end{aligned}$$

Read the last line back, symbol by symbol, into words: if the words are the sentence you began with, the translation is done.

Example. Five definitions, translated

Translate each sentence into a formula, stating the domain.

(a) $a_n \to L$: for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that $|a_n - L| < \varepsilon$ for every $n \geq N$.

(b) The set A of real numbers is bounded above.

(c) The function $f\colon \mathbb{R} \to \mathbb{R}$ is continuous at a: for every $\varepsilon > 0$ there is $\delta > 0$ such that $|f(x) - f(a)| < \varepsilon$ for every real x with $|x - a| < \delta$.

(d) There is no largest prime.

(e) Every integer greater than 1 has a prime divisor.

Solution. (a) Domain $\mathbb{R}$. The outermost claim is about every $\varepsilon > 0$: $\forall \varepsilon > 0$. What it says about ε is that some natural number N works: $\exists N \in \mathbb{N}$. What it says about N is that every natural number n from N on is good: $\forall n \in \mathbb{N}$, with the further condition $n \geq N$ under an implication. So

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to |a_n - L| < \varepsilon).$$

The phrase "for every $n \geq N$" carries two bounds on n, membership in $\mathbb{N}$ and $n \geq N$; the first is a bounded quantifier, the second the hypothesis of the implication. The free variable is L (and the statement is also about the function symbol a, which is not a variable).

(b) Domain $\mathbb{R}$; A names a set of reals, and $x \in A$ is a one-place predicate of x (Section 5.1). "Bounded above" means that some real number is at least every element of A. Outermost claim: there is an M. Its property: every element of A is at most M:

$$\exists M\ \forall x \in A\ (x \leq M).$$

The bound $x \in A$ belongs to the $\forall$; there is no bound on M, which may be any real number.

(c) Domain $\mathbb{R}$, with f a one-place function symbol and a a variable, free in the formula:

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x\ (|x - a| < \delta \to |f(x) - f(a)| < \varepsilon).$$

The last quantifier has no bound of the form "$x \in$ something", because f is defined on all of $\mathbb{R}$; its restriction $|x - a| < \delta$ goes under the implication, like $n \geq N$ in (a). For f defined on a set A, the last quantifier would be $\forall x \in A$.

(d) Domain $\mathbb{N}$, with $\text{Prime}(m)$ the predicate "m is prime". Section 5.1 read the sentence as $\neg\exists m\,(\text{Prime}(m) \wedge \forall n\,(\text{Prime}(n) \to n \leq m))$. A mathematician more often states it positively, "for every prime there is a larger prime":

$$\forall m\ (\text{Prime}(m) \to \exists n\ (\text{Prime}(n) \wedge m < n)).$$

The two forms say the same thing over $\mathbb{N}$: the laws of Section 6.2 push the $\neg$ inside and give $\forall m\,(\text{Prime}(m) \to \exists n\,(\text{Prime}(n) \wedge \neg(n \leq m)))$, and the order fact that $\neg(n \leq m)$ is the same as $m < n$ finishes the job. (That last step is a fact about the order of $\mathbb{N}$, not a law of logic.)

(e) Domain $\mathbb{Z}$. Outermost claim: every integer greater than 1 has a property, $\forall n > 1$. The property: some m is prime and divides n. "Divides" is itself a quantifier: $m \mid n$ means $n = m \cdot k$ for some integer k. So

$$\forall n > 1\ \exists m\ (\text{Prime}(m) \wedge \exists k\ (n = m \cdot k)).$$

Unfolded, the first bound is $\forall n\,(n > 1 \to \dots)$. The innermost quantifier came from a **defined** word, "divides", which is typical: a definition in words often hides quantifiers, and a translation unfolds them.

Two remarks on the example. First, the order of the quantifiers in (a) and (c) is the order in which the objects are **chosen**: ε first, by a challenger; then N or δ, by the defender, knowing ε; then n or x, by the challenger again, knowing both. That is the game reading of Example (Swapping two quantifiers). Second, in (c) the letter a is free, so the formula is a statement about a; "f is continuous at every point" would add $\forall a$ in front.

Quick check. Translate "A has a maximum" (an element of A that is at least every element of A) over $\mathbb{R}$. Which quantifier carries the bound $x \in A$, and does the other quantifier carry a bound too?

Solution. $\exists m \in A\, \forall x \in A\, (x \le m)$. The bound $x \in A$ belongs to the $\forall$: "every element of A". The $\exists$ carries a bound too, $m \in A$, because a maximum must itself lie in A; dropping it gives $\exists m\, \forall x \in A\, (x \le m)$, which is "$A$ is bounded above", a weaker statement. For $A = (0, 1)$ the weaker statement is true ($m = 1$) and the stronger one is false.

Warning. **Three mistranslations to avoid.** (1) An $\exists$ with $\to$ under it, such as $\exists x\,(x \in A \to x > 5)$: it is true as soon as something lies outside A (Section 5.1). (2) A bound on the wrong quantifier: "f is bounded on A" is $\exists M\, \forall x \in A\, (|f(x)| \le M)$; the form $\exists M \in A\, \forall x\, (|f(x)| \le M)$ asks for a bound that lies in A and works on all of $\mathbb{R}$, and it fails for $f(x) = x$ on $A = [0, 1]$, which is bounded on A by 1. (3) "For all large n" read as "for all n": "$a_n > 0$ for all large n" is $\exists N \in \mathbb{N}\, \forall n \in \mathbb{N}\, (n \ge N \to a_n > 0)$, not $\forall n \in \mathbb{N}\, (a_n > 0)$. The sequence $a_n = n - 3$ satisfies the first (take $N = 4$) and not the second ($a_0 = -3$).

6.1.3 Five traps

The warning above caught three slips. Here are five more, each with the reading that shows the wrong formula saying something else.

Example. Five traps, each with a separating reading

The domain is $\mathbb{R}$ unless the row says $\mathbb{N}$. A and B are sets of reals, and $x \in A$ is a one-place predicate (Section 5.1).

	Sentence	Wrong	Right	A reading that separates them
(a)	Some natural number lies strictly between $\frac{1}{3}$ and $\frac{2}{3}$.	$\exists n\,(\frac{1}{3} < n \wedge n < \frac{2}{3})$	$\exists n \in \mathbb{N}\,(\frac{1}{3} < n \wedge n < \frac{2}{3})$	The domain is $\mathbb{R}$, by step 1. The sentence is false, since $0 < \frac{1}{3}$ and $\frac{2}{3} < 1$; the wrong formula is true, with $n = \frac{1}{2}$. Under one domain, a letter for a natural number ranges over the reals unless its bound $n \in \mathbb{N}$ says otherwise.
(b)	Every element of A and of B is positive.	$\forall x\,((x \in A \wedge x \in B) \to x > 0)$	$\forall x\,((x \in A \vee x \in B) \to x > 0)$	$A = \{1\}$, $B = \{-1\}$: the sentence is false ($-1 \in B$). The wrong formula speaks only of common elements, and there are none, so it is true. The "and" of the sentence is an "or" in the restriction.
(c)	Only even numbers are multiples of 4. (Over $\mathbb{N}$.)	$\forall n\,(\text{Even}(n) \to 4 \mid n)$	$\forall n\,(4 \mid n \to \text{Even}(n))$	The right formula is true; the wrong one is false at $n = 2$. "Only E are M" puts E after the arrow, as "only if" did in Section 1.1.

(d)	If a natural number is divisible by 6, then it is even. (Over $\mathbb{N}$.)	$\exists n\,(6 \mid n) \to \text{Even}(n)$	$\forall n\,(6 \mid n \to \text{Even}(n))$	The right formula is true. In the wrong one the last n is free, and read at $n = 1$ it says "if some number is divisible by 6, then 1 is even": the hypothesis is true $(6 \mid 6)$ and the conclusion false. "A ... it" hides a "for every".
(e)	If any element of A is negative, then 0 is not a lower bound of A.	$\forall x \in A\,(x < 0) \to \neg\forall x \in A\,(0 \le x)$	$\exists x \in A\,(x < 0) \to \neg\forall x \in A\,(0 \le x)$	$A = \emptyset$: the right formula is true, since its hypothesis is false. In the wrong one the hypothesis "every element of $\emptyset$ is negative" is true and the conclusion "not every element of $\emptyset$ is at least 0" is false (Section 5.1, the empty set). This is Section 5.1's warning on "any", now with the reading that separates the two formulas: the empty set.

A reading like these is a counter-reading for the step from one formula to the other; it shows that the two say different things, which is all a translation needs to know.

Infinitely many, and eventually. "$a_n > 0$ for infinitely many n" says that however far out you start, some later term is positive:

$$\forall N \in \mathbb{N}\ \exists n \in \mathbb{N}\ (n \ge N \wedge a_n > 0).$$

"Eventually $a_n > 0$", that is, $a_n > 0$ for all large n (warning (3) above), says

$$\exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \ge N \to a_n > 0).$$

The quantifiers are swapped, and the bound $n \ge N$ moves from $\wedge$ to $\to$, as the two connectives of Definition 5.1.1 require. For $a_n = (-1)^n$ the first is true (every even $n \ge N$ will do) and the second is false (every odd $n \ge N$ fails).

Quick check. Translate "$a_n = 0$ for only finitely many n".

Solution. From some index on, no term is 0: $\exists N \in \mathbb{N}\, \forall n \in \mathbb{N}\, (n \geq N \to a_n \neq 0)$. Finitely many exceptions all lie below some N.

Translation error or derivation error? When a derivation ends in a statement you believe false, there are two places to look. A **derivation error** is an illegal line, and checking the derivation line by line finds it. A **translation error** is a formula that does not say what its sentence says, and no line check finds it, because every line may be correct. Suppose a reader translates "some element of A exceeds 5" as $\exists x\, (x \in A \to x > 5)$, the slip of warning (1) above, and takes $A = \{1, 2\}$. From the true line $\neg(7 \in A)$, two steps derive that formula: $7 \in A \to 7 > 5$ by Proposition 2.1.2, then $\exists$I at 7. Both steps are legal. The conclusion is a true formula about A, and the sentence it was meant to translate is false, since neither 1 nor 2 exceeds 5. The fault is in the translation, not in the derivation. So check a surprising conclusion in two ways: run the line check, and read the formulas back into words.

6.1.4 Rules for bounded quantifiers

Every analysis formula has bounds, and each bound doubles the work of a rule. To use $\forall \varepsilon > 0\, \varphi$ at the term 1, $\forall$E gives $1 > 0 \to \varphi[1/\varepsilon]$, and a second step, $\to$E with a line $1 > 0$, gives $\varphi[1/\varepsilon]$. To prove $\exists N \in \mathbb{N}\, \psi$ with the witness 7, $\wedge$I builds $7 \in \mathbb{N} \wedge \psi[7/N]$ and $\exists$I finishes. Every step is legal and every step is predictable, and a derivation about limits drowns in them. So we fold each pair into one step.

A bounded quantifier gets a rule of its own: the bound travels with the quantifier.

Proposition 6.1.1. Rules for Bounded Quantifiers

Let A be a fixed set, so that $\in A$ is a one-place predicate symbol (Section 5.2), x a variable and φ a formula. The following four rules, the **bounded rules**, may be used in derivations, with the rule names of the unbounded rules and the same side conditions: in $\forall x \in A$ E and $\exists x \in A$ I, the term t is free for x in φ; in $\forall x \in A$ I, the name x_0 appears in no premise, in no assumption of a box still open at line k, and not in $\forall x \in A\, \varphi$; in $\exists x \in A$ E, it appears in no premise, in no assumption of a box still open at line k, and not in χ or in $\exists x \in A\, \varphi$. The last line of each box is a line of that box itself.

i	$\forall x \in A\, \varphi$		i	$t \in A$	
j	$t \in A$		j	$\varphi[t/x]$	
k	$\varphi[t/x]$	$\forall$E i, j	k	$\exists x \in A\, \varphi$	$\exists$I i, j

					h	$\exists x \in A\ \varphi$		
i	x_0	$x_0 \in A$	asm, for $\forall$I		i	x_0 $x_0 \in A$	asm, for $\exists$E	
		$\vdots$			j	$\varphi[x_0/x]$	asm, for $\exists$E	
j		$\varphi[x_0/x]$				$\vdots$		
k	$\forall x \in A\ \varphi$		$\forall$I i–j		l	χ		
					k	χ	$\exists$E h, i–l	

Every use of a bounded rule can be replaced by a block of steps of the unbounded rules (Sections 5.3–5.6) and of Chapter 1, with the same premises, the same assumptions open, and the same last line. Hence any derivation that uses the bounded rules converts into one that does not. The same holds with $x \in A$ and $t \in A$ replaced by $x > a$ and $t > a$, or by the bounds $\geq a$, $< a$, $\leq a$, for a term a in which x does not occur.

In words: to **use** "every x in A has φ", produce an element t of A and conclude φ for t. To **prove** "some x in A has φ", exhibit an element t of A with φ. To **prove** "every x in A has φ", let x_0 be an arbitrary element of A (the box opens with the assumption $x_0 \in A$) and prove φ for it. To **use** "some x in A has φ", take such an element x_0, with both facts about it as assumptions, and derive something that does not mention it. These are exactly the sentences a mathematician writes: "let $\varepsilon > 0$", "since $\delta > 0$, the hypothesis applies to δ", "take $N = 7 \in \mathbb{N}$", "choose $\delta > 0$ such that ...".

The bounded $\forall$I box is the only flagged box with an assumption that closes by $\forall$I. Its assumption $x_0 \in A$ mentions the flag, and that is allowed: the side condition asks about assumptions of boxes **still open at line** k, and the box's own assumption has closed there.

In a finished derivation, the assumption $x_0 \in A$ of a bounded $\forall$I box is printed "asm, for $\forall$I", since the box closes by $\forall$I.

Second reading · *Why a proposition, not a definition*

Why is this a proposition and not simply a definition? Because the calculus is fixed: its rules are those of Chapters 1–2 and Sections 5.3–5.7, and a new rule is admissible only if it adds no derivable sequents. Definition 2.3.1 gave one way to show that: a proved sequent, written with statement letters, may be used as one line. It does not cover these rules. The bounded $\forall$E and $\exists$I are one line each, but they are schemas in a formula, a variable and a term, not in statement letters; the bounded $\forall$I and $\exists$E close boxes, which no proved sequent can. So the proposition proves admissibility the direct way: each bounded step is shown to be an abbreviation of a block of unbounded steps. In this sense the bounded rules are derived rules, in a looser sense that extends Chapter 2's: a line citing one abbreviates a block of primitive steps, not a proved sequent.

Idea. A bounded quantifier is an abbreviation (Definition 5.1.1): $\forall x \in A\ \varphi$ **is** the formula $\forall x\,(x \in A \to \varphi)$, and $\exists x \in A\ \varphi$ **is** $\exists x\,(x \in A \wedge \varphi)$. So each bounded rule should be an unbounded rule on the unfolded formula, plus the connective rule that handles $\to$ or $\wedge$. Write the block, then check two things: the block is a correct piece of derivation wherever

the bounded step was, and its side conditions hold exactly when the bounded ones do.

Warning (Not allowed: the bounded rules). Each step below is wrong; the tag names the condition of Proposition 6.1.1 that it breaks.

- From $\forall \varepsilon > 0\ \varphi$ alone, write $\varphi[1/\varepsilon]$: the bounded $\forall$E cites two lines, and the line $1 > 0$ is missing.
- From $|a_7 - L| < 1$ alone, write $\exists N \in \mathbb{N}\ (|a_N - L| < 1)$: the bounded $\exists$I needs the line $7 \in \mathbb{N}$ too.
- Close a bounded $\forall$I box on ε_0 while the assumption $\delta_0 < \varepsilon_0$ of an enclosing box is still open: the flag appears in an assumption of a box still open.
- Close a bounded $\exists$E box on N_0 with the last line $N_0 + 1 \in \mathbb{N}$: the conclusion χ mentions the witness.

The proof writes out, for each bounded rule, the block of unbounded steps it abbreviates, and checks that the side conditions match; it is in the fold below.

Second reading · *The proof: each bounded rule as a block of unbounded steps*

Proof. We give the replacement block for each rule and check it. In each block, the lines of the bounded version keep their labels, and new lines get new letters.

Bounded $\forall$E. Line i is $\forall x\,(x \in A \to \varphi)$. Replace line k by two lines:

i	$\forall x\,(x \in A \to \varphi)$	
j	$t \in A$	
m	$t \in A \to \varphi[t/x]$	$\forall$E i
k	$\varphi[t/x]$	$\to$E m, j

Line m is $\forall$E at t: the formula $(x \in A \to \varphi)[t/x]$ is $t \in A \to \varphi[t/x]$ by Definition 5.2.5, since A is not a variable. Its side condition asks that t be free for x in $x \in A \to \varphi$ (Definition 5.2.6), which holds exactly when t is free for x in φ, since $x \in A$ has no quantifier. Lines i and j are available at m and k because they were available at k.

Bounded $\exists$I. Replace line k by:

i	$t \in A$	
j	$\varphi[t/x]$	
m	$t \in A \wedge \varphi[t/x]$	$\wedge$I i, j
k	$\exists x\,(x \in A \wedge \varphi)$	$\exists$I m

Line m is the instance at t of $x \in A \wedge \varphi$, and t is free for x there exactly when it is free for x in φ.

Bounded $\forall$I. Replace the box and line k by a flagged box with no assumption, which contains

a →I box holding the old lines:

i	x_0	$x_0 \in A$		asm, for →I
		⋮		
j		$\varphi[x_0/x]$		
m		$x_0 \in A \to \varphi[x_0/x]$		→I i–j
k	$\forall x\,(x \in A \to \varphi)$			∀I i–m

The lines between i and j are unchanged. Each of them lies inside both new boxes, so every line it cited is still available (Definition 1.4.2, with the flagged boxes of Section 5.4): a line outside the old box is outside both new ones, and a line inside the old box is inside the new →I box. The side conditions inside the old box are unaffected too: a line inside it has the same open assumptions as before, $x_0 \in A$ (now the assumption of the →I box) together with the outer ones, because the flagged box adds none; so every flagged box that closes inside the old box still meets its side condition. Line m is →I, and it is the last line of the flagged box, as ∀I requires; it is the instance at x_0 of $x \in A \to \varphi$. Now the side condition of ∀I at line k: the premises are the same as before; the boxes still open at line k are the same as those open at line k in the bounded version, because the →I box on $x_0 \in A$ has **closed** before line k; and the conclusion is the same formula. So the unbounded side condition holds exactly when the bounded one does.

Bounded ∃E. Line h is $\exists x\,(x \in A \wedge \varphi)$. Replace the box's two assumptions by one assumption and two lines of ∧E:

h	$\exists x\,(x \in A \wedge \varphi)$			
g	x_0	$x_0 \in A \wedge \varphi[x_0/x]$		asm, for ∃E
i		$x_0 \in A$		∧E g
j		$\varphi[x_0/x]$		∧E g
		⋮		
l		χ		
k	χ			∃E h, g–l

Lines i and j now hold the same formulas as the old assumptions, inside the same box, so every line that cited them may still cite them. At a line inside the box, the open assumptions are now g together with the outer ones, where before they were $x_0 \in A$ and $\varphi[x_0/x]$ together with the outer ones. A name appears in g exactly when it appears in $x_0 \in A$ or in $\varphi[x_0/x]$, so every flagged box that closes inside the box still meets its side condition. The box assumes the instance at x_0 of $x \in A \wedge \varphi$, as ∃E requires. Its side condition at line k names χ, the formula $\exists x\,(x \in A \wedge \varphi)$, the premises and the assumptions open at line k: the same four things as the bounded side condition, with the same result.

For the bounds $x > a$ and the others, the same four blocks work with $x > a$ and $t > a$ in place of $x \in A$ and $t \in A$. One point needs the hypothesis that x does not occur in a: then $(x > a)[t/x]$ is $t > a$, and the bound contains no bound variable, so free-for conditions are unaffected by it.

Replacing the bounded steps one at a time, each by its block, removes them all; the premises

and the last line do not change. This proves the proposition.

■

Each block is one line longer than the bounded step it replaces: two lines for one in ∀E and ∃I; the extra →I line in ∀I; three lines (an assumption and two ∧E) for two assumptions in ∃E. So **each bounded step saves exactly one line**, and a derivation with many bounds saves many. More than the count, the bounded rules keep the derivation at the level at which a mathematician thinks: "let $\varepsilon > 0$" is one step, not two.

Quick check. Unfolded, $\exists N \in \mathbb{N}\ \psi$ is $\exists N\,(N \in \mathbb{N} \wedge \psi)$. A student uses the bounded ∃E on it and writes the box with the single assumption $N_0 \in \mathbb{N} \wedge \psi[N_0/N]$. Is that wrong?

Solution. No, it is the unbounded ∃E, which is always available, since the bounded formula **is** the unfolded one. It costs two ∧E lines inside the box to separate $N_0 \in \mathbb{N}$ from $\psi[N_0/N]$. The bounded rule opens the box with the two assumptions already separated. Both are correct; the bounded one is shorter by one line.

6.1.5 A limit derivation with bounds

Here is the first derivation in the book about a limit. The claim is the school fact "if $a_n \to L$, then $2a_n \to 2L$", with the definition of Example (Five definitions, translated) (a).

Claim. Over $\mathbb{R}$,

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to |a_n - L| < \varepsilon)$$

derives the same formula with $2 \cdot a_n$ and $2 \cdot L$ in place of a_n and L, using two lines of arithmetic, justified "arith": $\varepsilon_0 > 0$ gives $\varepsilon_0/2 > 0$, and $|a_{n_0} - L| < \varepsilon_0/2$ gives $|2 \cdot a_{n_0} - 2 \cdot L| < \varepsilon_0$ (because $|2u - 2v| = 2|u - v|$).

Idea. Run the search. The goal starts with $\forall \varepsilon > 0$: open a bounded ∀I box, "let $\varepsilon_0 > 0$". The new goal starts with $\exists N \in \mathbb{N}$: **postpone** it (the step of Section 5.5), since no witness is in sight. Turn to the premise: it starts with $\forall \varepsilon > 0$, and the step of Section 5.3 says to instantiate it at the term the goal needs. Which term? The goal will need $|2a_n - 2L| < \varepsilon_0$, which follows from $|a_n - L| < \varepsilon_0/2$: so instantiate at $\varepsilon_0/2$, after an "arith" line $\varepsilon_0/2 > 0$ for the bound. That gives an $\exists N \in \mathbb{N}$ line: open its ∃E box **at once**, with a witness N_0 (the step of Section 5.6), and let the box's last line be the current goal. Now the postponed goal has its witness: N_0 itself. The remaining goal $\forall n \in \mathbb{N}\,(\dots)$ opens a bounded ∀I box for n_0, and a →I box on $n_0 \geq N_0$ finishes with one ∀E, one →E, and the second "arith" line. This is the pattern of Proposition 5.6.2: **fix** ε_0, **obtain** N_0, **choose** N_0 as the witness.

Two abbreviations show the shape. Write $G(N, \varepsilon)$ for $\forall n \in \mathbb{N}\ (n \geq N \to |a_n - L| < \varepsilon)$, "from N on, every term is within ε of L", and $G_2(N, \varepsilon)$ for the same formula with $2 \cdot a_n$ and $2 \cdot L$. The premise is $\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ G(N, \varepsilon)$, and the goal is the same with G_2. With the lines that look inside G left as a gap, the derivation is:

1		$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ G(N, \varepsilon)$	
2	ε_0	$\varepsilon_0 > 0$	asm, for $\forall$I
3		$\varepsilon_0/2 > 0$	arith 2
4		$\exists N \in \mathbb{N}\ G(N, \varepsilon_0/2)$	$\forall$E 1, 3
5	N_0	$N_0 \in \mathbb{N}$	asm, for $\exists$E
6		$G(N_0, \varepsilon_0/2)$	asm, for $\exists$E
		$\vdots$	
13		$G_2(N_0, \varepsilon_0)$	
14		$\exists N \in \mathbb{N}\ G_2(N, \varepsilon_0)$	$\exists$I 5, 13
15		$\exists N \in \mathbb{N}\ G_2(N, \varepsilon_0)$	$\exists$E 4, 5–14
16		$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ G_2(N, \varepsilon)$	$\forall$I 2–15

Lines 7–12 are the only lines that look inside G and G_2: they unfold $G(N_0, \varepsilon_0/2)$ at a fresh n_0 and prove the instance that line 13 folds back into $G_2(N_0, \varepsilon_0)$. The derivation in full follows the check below; its lines 1–6 and 13–16 are the lines above, written out.

Quick check. In the outline, line 4 instantiates the premise at $\varepsilon_0/2$. Why that term, and not ε_0?

Solution. The goal needs $|2a_n - 2L| < \varepsilon_0$, which is $2|a_n - L| < \varepsilon_0$, so it is enough to have $|a_n - L| < \varepsilon_0/2$. The term is read off the goal (Section 5.3's step), not off the premise. At ε_0 the premise would give only $|a_n - L| < \varepsilon_0$, hence $|2a_n - 2L| < 2\varepsilon_0$, which does not reach the goal.

1		$\forall \varepsilon > 0\, \exists N \in \mathbb{N}\, \forall n \in \mathbb{N}\, (n \geq N \to \lvert a_n - L \rvert < \varepsilon)$	
2	ε_0	$\varepsilon_0 > 0$	asm, for $\forall$I
3		$\varepsilon_0/2 > 0$	arith 2
4		$\exists N \in \mathbb{N}\, \forall n \in \mathbb{N}\, (n \geq N \to \lvert a_n - L \rvert < \varepsilon_0/2)$	$\forall$E 1, 3
5	N_0	$N_0 \in \mathbb{N}$	asm, for $\exists$E
6		$\forall n \in \mathbb{N}\, (n \geq N_0 \to \lvert a_n - L \rvert < \varepsilon_0/2)$	asm, for $\exists$E
7	n_0	$n_0 \in \mathbb{N}$	asm, for $\forall$I
8		$n_0 \geq N_0 \to \lvert a_{n_0} - L \rvert < \varepsilon_0/2$	$\forall$E 6, 7
9		$n_0 \geq N_0$	asm, for $\to$I
10		$\lvert a_{n_0} - L \rvert < \varepsilon_0/2$	$\to$E 8, 9
11		$\lvert 2 \cdot a_{n_0} - 2 \cdot L \rvert < \varepsilon_0$	arith 10
12		$n_0 \geq N_0 \to \lvert 2 \cdot a_{n_0} - 2 \cdot L \rvert < \varepsilon_0$	$\to$I 9–11
13		$\forall n \in \mathbb{N}\, (n \geq N_0 \to \lvert 2 \cdot a_n - 2 \cdot L \rvert < \varepsilon_0)$	$\forall$I 7–12
14		$\exists N \in \mathbb{N}\, \forall n \in \mathbb{N}\, (n \geq N \to \lvert 2 \cdot a_n - 2 \cdot L \rvert < \varepsilon_0)$	$\exists$I 5, 13
15		$\exists N \in \mathbb{N}\, \forall n \in \mathbb{N}\, (n \geq N \to \lvert 2 \cdot a_n - 2 \cdot L \rvert < \varepsilon_0)$	$\exists$E 4, 5–14
16		$\forall \varepsilon > 0\, \exists N \in \mathbb{N}\, \forall n \in \mathbb{N}\, (n \geq N \to \lvert 2 \cdot a_n - 2 \cdot L \rvert < \varepsilon)$	$\forall$I 2–15

Check the side conditions at the three lines that close flagged boxes, and the witness line.

- Line 13 ($\forall n \in \mathbb{N}$ I, flag n_0): n_0 is not in the premise, not in line 13, and not in the assumptions still open at line 13, which are lines 2, 5, and 6.
- Line 14 ($\exists N \in \mathbb{N}$ I): the witness N_0, with its bound on line 5 and the instance on line 13.
- Line 15 ($\exists N \in \mathbb{N}$ E, flag N_0): N_0 is not in line 15, the conclusion (the witness has been forgotten by line 14), not in line 4, not in the premise, and not in the open assumption on line 2.
- Line 16 ($\forall \varepsilon > 0$ I, flag ε_0): ε_0 is not in the premise and not in line 16, and no assumption is open at line 16.

Line 4 is where N comes to depend on ε: the witness N_0 is obtained from the premise **at** $\varepsilon_0/2$, so it depends on ε_0, exactly as the order $\forall \varepsilon\, \exists N$ allows. Section 6.4 makes this reading of dependence precise.

In prose, line by line:

> Let $\varepsilon > 0$ [lines 2, 16]. Then $\varepsilon/2 > 0$ [3], so, since $a_n \to L$, there is $N \in \mathbb{N}$ with $\lvert a_n - L \rvert < \varepsilon/2$ for every $n \geq N$ [4]; choose such an N [5, 6, 15]. Let $n \in \mathbb{N}$ with $n \geq N$ [7, 9]. Then $\lvert a_n - L \rvert < \varepsilon/2$ [8, 10], so $\lvert 2a_n - 2L \rvert = 2\lvert a_n - L \rvert < \varepsilon$ [11]. So this N works [12–14].

Written with the unbounded rules, the same derivation has 22 lines, since each of its six bounded steps (lines 4, 8, 13, 14, 15, and 16) costs one line more. The claim "$\frac{1}{n} \to 0$" (for $n \geq 1$) is not derived here: its proof needs the Archimedean property, a fact about $\mathbb{N}$ inside $\mathbb{R}$, not a law of logic.

Remark (Looking ahead). Induction ("a set of natural numbers that contains 0 and contains $n+1$ whenever it contains n contains every natural number") and the least-upper-bound property of the real numbers say "for every **set** of numbers", and the variables of this chapter range over numbers only. Chapter 7's optional last section says what that costs.

6.1.6 Exercises

6.1.6.1 A. Check your understanding

Exercise. A1

(a) Translate into formulas, stating the domain: (i) every real number has a cube root; (ii) the function f is bounded on A; (iii) between any two different real numbers there is a rational number; (iv) the sequence (a_n) is eventually constant; (v) some natural number divides every natural number.

(b) Three students translate "some element of A is larger than every element of B", over $\mathbb{R}$:
   (i) $\exists x \in A\ \forall y \in B\ (x > y)$; (ii) $\exists x\,(x \in A \to \forall y \in B\ (x > y))$; (iii) $\forall y \in B\ \exists x \in A\ (x > y)$. Which are correct? For each wrong one, give sets A, B for which it has the wrong truth value.

(c) Which variables are free in your answer to (a)(iv)?

(d) Write the bounded rule $\exists x \in A$ I as a block of two unbounded steps, and name the two rules.

Solution. *(a)* (i) Over $\mathbb{R}$: $\forall x\,\exists y\,(y \cdot y \cdot y = x)$. (ii) Over $\mathbb{R}$: $\exists M\,\forall x \in A\,(|f(x)| \le M)$. (iii) Over $\mathbb{R}$, with $\in \mathbb{Q}$ a one-place predicate symbol: $\forall x\,\forall y\,(x < y \to \exists q \in \mathbb{Q}\ (x < q \wedge q < y))$; "different" is carried by $x < y$, which names the smaller number x (the case $y < x$ is the same statement with the letters exchanged). (iv) Over $\mathbb{R}$, with a a function symbol: $\exists N \in \mathbb{N}\,\forall n \in \mathbb{N}\,(n \ge N \to a_n = a_N)$. (v) Over $\mathbb{N}$: $\exists d\,\forall n\,\exists k\,(n = d \cdot k)$; it is true, with $d = 1$.

(b) Only (i) is correct. (ii) has $\to$ under $\exists$: with $A = \{0\}$ and $B = \{5\}$, the sentence in words is false, but (ii) is true, because $x = 7 \notin A$ makes the implication vacuously true. (iii) has the quantifiers in the wrong order: it lets the element of A depend on y. With $A = B = \mathbb{N}$, (iii) is true (for each y, take $x = y + 1$), and the sentence in words is false, since no natural number exceeds every natural number.

(c) None of the variables: N and n are bound. The formula is a statement about the function symbol a, which is not a variable.

(d) From $t \in A$ and $\varphi[t/x]$: first $t \in A \wedge \varphi[t/x]$ by $\wedge$I, then $\exists x\,(x \in A \wedge \varphi)$ by $\exists$I, which is $\exists x \in A\ \varphi$ unfolded.

Exercise. A2: Drill

Each part takes a line or two. The domain is $\mathbb{R}$, and A, B are sets of reals. The parts run from translating, through unfolding and judging, to single steps of the bounded rules.

(a) Translate "f is bounded above on A".

(b) Translate "every element of A is less than some element of B".

(c) Translate "some element of A is less than every element of B".

(d) Unfold the bounds of (c).

(e) Write the mixed stage, then the formula, for "every element of A has a larger element in A".

(f) Translate "$a_n > 0$ for all large n".

(g) Translate "$a_n \leq 0$ for infinitely many n".

(h) Translate "not every element of A is positive" and "no element of A is positive", and decide both for $A = \{-1, 1\}$.

(i) Translate "if $x \cdot x < 0$ for any real x, then $0 = 1$".

(j) Lines 1 $\forall\varepsilon > 0\ \varphi$ and 2 $1 > 0$. What does the bounded $\forall$E 1, 2 give?

(k) Lines 1 $3 \in \mathbb{N}$ and 2 $3+3 = 6$. What does the bounded $\exists$I 1, 2 give, with φ the formula $n + n = 6$?

(l) How many lines does a bounded $\exists$E save over the unbounded rule?

Solution.

(a) $\exists M\, \forall x \in A\ (f(x) \leq M)$.

(b) $\forall x \in A\, \exists y \in B\ (x < y)$.

(c) $\exists x \in A\, \forall y \in B\ (x < y)$.

(d) $\exists x\, (x \in A \wedge \forall y\, (y \in B \to x < y))$.

(e) $\forall x \in A$ [some element of A is larger than x], then $\forall x \in A\, \exists y \in A\ (x < y)$.

(f) $\exists N \in \mathbb{N}\, \forall n \in \mathbb{N}\ (n \geq N \to a_n > 0)$.

(g) $\forall N \in \mathbb{N}\, \exists n \in \mathbb{N}\ (n \geq N \wedge a_n \leq 0)$.

(h) $\neg\forall x \in A\ (x > 0)$, true for $\{-1, 1\}$ (because of -1); $\neg\exists x \in A\ (x > 0)$, or $\forall x \in A\ \neg(x > 0)$, false (because of 1). One word apart, opposite truth values.

(i) $\exists x\, (x \cdot x < 0) \to 0 = 1$: "any" after "if" is "some". It is true, since no real square is negative.

(j) $\varphi[1/\varepsilon]$.

(k) $\exists n \in \mathbb{N}\ (n + n = 6)$.

(l) One: two assumptions in place of one assumption and two $\wedge$E lines.

6.1.6.2 B. Practice

Exercise. B1: Three definitions

Formalize each statement, stating the domain and the kind of each symbol (variable, constant, function symbol, predicate symbol; membership $\in A$ in a fixed set A is a one-place predicate symbol, as in Section 5.2).

(a) The set A of real numbers is bounded (above and below).

(b) The function $f\colon \mathbb{R} \to \mathbb{R}$ is increasing on A: for all $x, y \in A$, $x < y$ implies $f(x) \leq f(y)$.

(c) The sequence (a_n) is bounded: there is a real M with $|a_n| \leq M$ for all n.

Then, for (c), write the formula with the two quantifiers in the other order, and say why it

holds for every sequence.

Solution. No constants occur in any of the three parts.

(a) Domain $\mathbb{R}$; m, M, x variables (all bound); $\in A$ a one-place and $\leq$ a two-place predicate symbol. $\exists m\, \exists M\, \forall x \in A\, (m \leq x \wedge x \leq M)$. The two witnesses are chosen before x, so they must work for every element of A.

(b) Domain $\mathbb{R}$; x, y variables (bound); f a one-place function symbol; $\in A$ a one-place and $<$, $\leq$ two-place predicate symbols. $\forall x \in A\, \forall y \in A\, (x < y \to f(x) \leq f(y))$.

(c) Domain $\mathbb{R}$; M, n variables (bound); a a one-place function symbol (the one-domain convention, step 1 of the procedure), and $|\cdot|$ a one-place function symbol; $\in \mathbb{N}$ a one-place and $\leq$ a two-place predicate symbol. $\exists M\, \forall n \in \mathbb{N}\, (|a_n| \leq M)$.

Swapped: $\forall n \in \mathbb{N}\, \exists M\, (|a_n| \leq M)$. It holds for every sequence, because the witness may depend on n: take $M = |a_n|$. The sequence $a_n = n$, which is not bounded, satisfies it. As in Example (Swapping two quantifiers), the swap turns a statement about one M for all n into a statement about each term separately.

Exercise. B2: A bounded rehearsal

Using the bounded rules of Proposition 6.1.1, derive

$$\forall x \in A\, (P(x) \to Q(x)),\ \ \exists x \in A\, P(x)\ \vdash\ \exists x \in A\, Q(x).$$

Then say how many lines the derivation would have with the unbounded rules only.

Solution. The goal is an $\exists$: postpone it. The second premise is an $\exists$: open its bounded $\exists$E box at once, with a witness x_0 and the two assumptions $x_0 \in A$ and $P(x_0)$, and aim for the goal as the box's last line. Inside, the first premise is a $\forall x \in A$, and the term the box offers is x_0, with its bound on a line. The witness for the goal is x_0 again.

1		$\forall x \in A\, (P(x) \to Q(x))$	
2		$\exists x \in A\, P(x)$	
3	x_0	$x_0 \in A$	asm, for $\exists$E
4		$P(x_0)$	asm, for $\exists$E
5		$P(x_0) \to Q(x_0)$	$\forall$E 1, 3
6		$Q(x_0)$	$\to$E 5, 4
7		$\exists x \in A\, Q(x)$	$\exists$I 3, 6
8		$\exists x \in A\, Q(x)$	$\exists$E 2, 3–7

At line 8, x_0 is not in line 8, not in line 2, and not in either premise; no other box is open. The derivation uses three bounded steps (lines 5, 7 and 8), each of which costs one more line unbounded, so the unbounded derivation has $8 + 3 = 11$ lines: the $\exists$E box assumes $x_0 \in A \wedge P(x_0)$ and splits it with two $\wedge$E lines; line 5 becomes $\forall$E then $\to$E; line 7 becomes

∧I then ∃I.

The bound is the whole point of the exercise: the witness x_0 must carry its membership in A from the premise to the conclusion, and the bounded rules make that one assumption and one citation.

Exercise. B3: One key, twenty sentences

The domain is $\mathbb{N} = \{0, 1, 2, \dots\}$. Use only $\text{Prime}(x)$ ("x is prime"), $\text{Even}(x)$, $x < y$, $x \mid y$ ("x divides y", that is, $y = x \cdot k$ for some k), $=$ and the constants 1, 2, 9, with variables x, y, z. Translate sentences 1–18 into formulas. Then translate 19 and 20 into English and decide whether each is true.

1. 2 is an even prime.
2. Some prime divides 9.
3. Every prime greater than 2 is odd.
4. No prime is less than 2.
5. Every even number is divisible by 2.
6. Only odd numbers divide 9.
7. 1 divides every number.
8. Some number divides every number.
9. Every number is divisible by some number.
10. There is no largest prime.
11. There are infinitely many primes.
12. Every number greater than 1 has a prime divisor.
13. The only divisors of a prime are 1 and the prime itself.
14. If a number divides 2, it is 1 or 2.
15. Not every odd number is prime.
16. Between any two numbers, one less than the other, there is a third.
17. Some prime divides every even number.
18. Every divisor of 1 is 1.
19. $\forall x \, \forall y \, ((x \mid y \wedge y \mid x) \to x = y)$.
20. $\forall x \, (\text{Even}(x) \to \exists y \, (\text{Prime}(y) \wedge y \mid x \wedge \neg(y = 2)))$.

Solution.

1. $\text{Even}(2) \wedge \text{Prime}(2)$.
2. $\exists x \, (\text{Prime}(x) \wedge x \mid 9)$.
3. $\forall x \, ((\text{Prime}(x) \wedge 2 < x) \to \neg\text{Even}(x))$.
4. $\neg\exists x \, (\text{Prime}(x) \wedge x < 2)$, or $\forall x \, (\text{Prime}(x) \to \neg(x < 2))$.
5. $\forall x \, (\text{Even}(x) \to 2 \mid x)$.
6. $\forall x \, (x \mid 9 \to \neg\text{Even}(x))$.
7. $\forall x \, (1 \mid x)$.
8. $\exists x \, \forall y \, (x \mid y)$.
9. $\forall y \, \exists x \, (x \mid y)$.
10. $\forall x \, (\text{Prime}(x) \to \exists y \, (\text{Prime}(y) \wedge x < y))$.
11. $\forall x \, \exists y \, (x < y \wedge \text{Prime}(y))$.
12. $\forall x \, (1 < x \to \exists y \, (\text{Prime}(y) \wedge y \mid x))$.
13. $\forall x \, (\text{Prime}(x) \to \forall y \, (y \mid x \to (y = 1 \vee y = x)))$.
14. $\forall x \, (x \mid 2 \to (x = 1 \vee x = 2))$.

15. $\neg\forall x\,(\neg\text{Even}(x) \to \text{Prime}(x))$.
16. $\forall x\,\forall y\,(x < y \to \exists z\,(x < z \wedge z < y))$.
17. $\exists x\,(\text{Prime}(x) \wedge \forall y\,(\text{Even}(y) \to x \mid y))$.
18. $\forall x\,(x \mid 1 \to x = 1)$.
19. "Two numbers that divide each other are equal." True in $\mathbb{N}$: if $y = xk$ and $x = yl$, then $x = xkl$, so either $x = 0$, and then $y = 0$ too, or $kl = 1$, which in $\mathbb{N}$ means $k = l = 1$.
20. "Every even number has a prime divisor other than 2." False: at $x = 2$, and at every power of 2 $(4, 8, 16, 32, \ldots)$, the only prime divisor is 2.

Sentence 10 restricts x to primes and 11 does not; both are true, and in $\mathbb{N}$ each follows from the other. Sentence 16 is false (take $x = 0$, $y = 1$), and its translation is still correct: a correct translation of a false sentence is a false formula.

Exercise. B4: Unfold, then fold back

Over $\mathbb{R}$, for fixed sets A, B, C of reals, write $A \subseteq B$ for "every element of A is an element of B", $A \cup B$ for the set of reals in A or in B, and $A \cap B$ for the set of those in both. Use only the predicates $\in A$, $\in B$, $\in C$.

(a) Unfold $A \subseteq B \cap C$.

(b) Unfold $A \cup B \subseteq C$.

(c) Unfold "$A \cap B$ has no element".

(d) Fold back $\forall x\,(x \in A \to x \in B) \wedge \forall x\,(x \in B \to x \in C)$.

(e) Fold back $\forall x\,((x \in A \wedge x \in B) \to x \in C)$.

(f) A reader unfolds $A \cup B \subseteq C$ as the formula in (e). Give sets that separate the two, and say why the reader's formula is wrong.

Solution. *(a)* $\forall x\,(x \in A \to (x \in B \wedge x \in C))$.

(b) $\forall x\,((x \in A \vee x \in B) \to x \in C)$.

(c) $\neg\exists x\,(x \in A \wedge x \in B)$.

(d) $A \subseteq B$ and $B \subseteq C$.

(e) $A \cap B \subseteq C$.

(f) $A = \{1\}$, $B = \{2\}$, $C = \emptyset$. Then $A \cup B \subseteq C$ is false, since $1 \in A \cup B$ and $1 \notin C$. The formula of (e) is true, because $A \cap B$ has no element, so its hypothesis $x \in A \wedge x \in B$ is false for every x. The reader's mistake is trap (b) of the section: "in A or in B" is the restriction of (b), and it must be written with $\vee$; the "and" of "every element of A and of B" is an "or" in the restriction.

Marking scheme. (10 points) - 2: (a)–(c) together, with every bound attached by $\to$ under $\forall$ (and by $\wedge$ under the $\exists$ of (c)). - 2: (d) and (e). - 3: (f): the two sets, with one sentence on each truth value. - 3: the reason in words that the restriction of (b) is an "or". Fatal error (0 for (f)): a pair of sets on which the two formulas agree.

6.1.6.3 C. Going deeper

Exercise. C1: Where the $\forall x$ stands

Let $f\colon \mathbb{R} \to \mathbb{R}$ and $A \subseteq \mathbb{R}$. Translate, over $\mathbb{R}$:

(a) f is **uniformly continuous** on A: for every $\varepsilon > 0$ there is $\delta > 0$ such that for all $x, y \in A$, $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$;

(b) f is continuous at every point of A: for every $x \in A$ and every $\varepsilon > 0$ there is $\delta > 0$ such that for every $y \in A$, $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$.

Then describe, in one sentence, the only difference between the two formulas, and say what it means for the choice of δ.

Solution. (a) $\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A\ \forall y \in A\ \big(|x - y| < \delta \to |f(x) - f(y)| < \varepsilon\big)$.

(b) $\forall x \in A\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall y \in A\ \big(|x - y| < \delta \to |f(x) - f(y)| < \varepsilon\big)$.

The only difference is the position of $\forall x \in A$: in (a) it stands after $\exists \delta > 0$, in (b) before it. In (b) the defender chooses δ knowing x, so δ may depend on x and on ε; in (a) δ is chosen before any x is given, so one δ must serve every $x \in A$ at once, and it depends on ε only. Section 6.4 derives (b) from (a) and shows that x^2 on $\mathbb{R}$ satisfies (b) but not (a); Section 7.2 turns this example into a counter-model, so (b) does not derive (a).

Exercise. C2: Six orders, four statements

Fix a formula μ (the "matrix") in the variables x, y, z, and consider the six sentences obtained by putting the prefixes $\forall x$, $\exists y$, $\forall z$ in front of μ in the six possible orders: $\forall x\, \exists y\, \forall z\, \mu$, $\forall z\, \exists y\, \forall x\, \mu$, and so on. Call two of these orders **essentially the same** if the two sentences are provably equivalent ($\dashv\vdash$) for every matrix μ, and **essentially different** if some matrix over $\mathbb{Z}$ gives them different truth values (read informally, as in Section 5.1).

(a) Show that the six orders fall into four groups, with any two orders in one group essentially the same.

(b) Using the matrices $y > x$ and $y > z$ over $\mathbb{Z}$, show that any two orders from different groups are essentially different. So the six orders say four different things.

Solution (Hint 1). Two $\forall$'s side by side can be exchanged: that was one of the derivations of Example (Flagged boxes at work). What is left to decide is where $\exists y$ stands.

Solution (Hint 2). List the six orders by the position of $\exists y$: first, middle or last. When it is first or last, the two $\forall$'s are side by side.

Solution. (a) Sort the orders by the position of $\exists y$.

- $\exists y$ first: $\exists y\, \forall x\, \forall z\, \mu$ and $\exists y\, \forall z\, \forall x\, \mu$.
- $\exists y$ last: $\forall x\, \forall z\, \exists y\, \mu$ and $\forall z\, \forall x\, \exists y\, \mu$.
- $\exists y$ in the middle: $\forall x\, \exists y\, \forall z\, \mu$, and $\forall z\, \exists y\, \forall x\, \mu$.

Two adjacent ∀'s may be exchanged: $\forall x\,\forall z\,\psi \dashv\vdash \forall z\,\forall x\,\psi$ for every formula ψ, by the derivation of Example (Flagged boxes at work) that exchanges two universal quantifiers (it uses nothing about its matrix but that it is a formula). This gives the "last" pair at once. For the "first" pair the ∀'s sit under $\exists y$: open the ∃E box on $\exists y\,\forall x\,\forall z\,\mu$ with a witness y_0, exchange the two ∀'s inside it, and close with ∃I at y_0; the other direction is the same. So there are at most four groups: $\{\exists y\,\forall x\,\forall z,\ \exists y\,\forall z\,\forall x\}$, $\{\forall x\,\forall z\,\exists y,\ \forall z\,\forall x\,\exists y\}$, $\{\forall x\,\exists y\,\forall z\}$ and $\{\forall z\,\exists y\,\forall x\}$.

(b) Call the groups F ($\exists y$ first), L (last), M_x ($\forall x\,\exists y\,\forall z$) and M_z ($\forall z\,\exists y\,\forall x$). Over $\mathbb{Z}$, evaluate each group with the two matrices, reading informally (Section 5.1):

Group	$\mu = (y > x)$	$\mu = (y > z)$
F	false: no y exceeds every x	false: no y exceeds every z
L	true: given x, z, take $y = x + 1$	true: take $y = z + 1$
M_x	true: given x, take $y = x + 1$	false: no y exceeds every z
M_z	false: no y exceeds every x	true: given z, take $y = z + 1$

Every two rows differ in at least one column, so every two groups are separated by one of the two matrices: F from L, from M_x and from M_z by the column in which the other group has "true"; L from M_x by $y > z$ and from M_z by $y > x$; M_x from M_z by either. So the six orders say four different things: within a group the sentences are provably equivalent by (a), and across groups some matrix over $\mathbb{Z}$ tells them apart.

Marking scheme. (10 points) - 2: lists the six orders and sorts them by the position of $\exists y$. - 3: justifies the two collapses, citing the exchange of adjacent ∀'s, and handles the pair under $\exists y$ (by an ∃E box, or equivalently). - 3: gives, for each of the six pairs of groups, a matrix and the two truth values, with witnesses for the true sentences. - 2: concludes that the six orders say four different things, combining (a) and (b). Fatal errors (0 for the step): exchanging ∀ with ∃; a "separation" in which the two sentences have the same truth value.

Exercise. C3: Eventually, therefore infinitely often

(a) With the bounded rules, derive

$$\exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to P(n)) \ \vdash\ \forall M \in \mathbb{N}\ \exists n \in \mathbb{N}\ (n \geq M \wedge P(n)).$$

Lines of arithmetic may be justified "arith".

(b) For the converse, give a reading of P over $\mathbb{N}$ under which the premise is true and the conclusion false.

Solution. (a) The goal is a bounded ∀: open its box first, with $M_0 \in \mathbb{N}$. The premise is a bounded ∃: open its box early, with $N_0 \in \mathbb{N}$ and its property. The postponed goal $\exists n \in \mathbb{N}$ needs a term that is at least M_0 and at which the premise applies, so at least N_0: the sum $N_0 + M_0$ is both.

1		$\exists N \in \mathbb{N}\, \forall n \in \mathbb{N}\, (n \geq N \to P(n))$	
2	M_0	$M_0 \in \mathbb{N}$	asm, for $\forall$I
3	N_0	$N_0 \in \mathbb{N}$	asm, for $\exists$E
4		$\forall n \in \mathbb{N}\, (n \geq N_0 \to P(n))$	asm, for $\exists$E
5		$N_0 + M_0 \in \mathbb{N}$	arith 2, 3
6		$N_0 + M_0 \geq N_0 \to P(N_0 + M_0)$	$\forall$E 4, 5
7		$N_0 + M_0 \geq N_0$	arith 2
8		$P(N_0 + M_0)$	$\to$E 6, 7
9		$N_0 + M_0 \geq M_0$	arith 3
10		$N_0 + M_0 \geq M_0 \wedge P(N_0 + M_0)$	$\wedge$I 9, 8
11		$\exists n \in \mathbb{N}\, (n \geq M_0 \wedge P(n))$	$\exists$I 5, 10
12		$\exists n \in \mathbb{N}\, (n \geq M_0 \wedge P(n))$	$\exists$E 1, 3–11
13		$\forall M \in \mathbb{N}\, \exists n \in \mathbb{N}\, (n \geq M \wedge P(n))$	$\forall$I 2–12

Lines 7 and 9 use $M_0 \geq 0$ and $N_0 \geq 0$, which hold for natural numbers. The side conditions:

- At line 12, N_0 is not in line 12, not in line 1, not in the premise, and not in the open assumption on line 2.
- At line 13, M_0 is not in the premise and not in line 13, and no assumption is open.
- At lines 6 and 11, the term $N_0 + M_0$ is free for n.

The witness $N_0 + M_0$ is built from the two names in play: fix M_0, obtain N_0, choose a term made from both.

(b) Let $P(n)$ be "n is even". The premise of the converse, $\forall M \in \mathbb{N}\, \exists n \in \mathbb{N}\, (n \geq M \wedge P(n))$, is true: beyond any M there is an even number, such as $2M$. Its conclusion, $\exists N \in \mathbb{N}\, \forall n \in \mathbb{N}\, (n \geq N \to P(n))$, is false: beyond any N there is an odd number, $2N + 1$. This is a counter-reading (Section 7.2).

Marking scheme. (10 points) - 2: opens the flagged box of M_0 and the $\exists$E box on the premise (in either nesting) before choosing the witness. - 3: a witness that depends on both names, with both bounds checked by arith lines. - 2: checks the side conditions at lines 12 and 13. - 3: (b): a reading, with both truth values justified. Fatal errors (0 for the step): a witness chosen before M_0 is given; the witness N_0 itself, which need not be at least M_0.

6.2 Negating and Moving Quantifiers

To prove that a sequence does **not** converge, or that a function is **not** uniformly continuous, we must use the negation of a definition, and a negation is only usable once the $\neg$ has been pushed all the way in. "It is not the case that for every $\varepsilon > 0$ there is an $N \ldots$" gives a proof nothing to hold. "There is an $\varepsilon > 0$ such that for every N there is an $n \geq N$ with $\ldots$" tells it what to find. Section 5.1 read the step from the first form to the second informally, as the challenger and the defender changing sides. This section derives the laws that make it exact, with the rules of Sections 5.3–5.6. Once they are proved, negating a definition is mechanical: one law

per symbol, from the outside in. It then shows the laws at work, in an analysis proof in which a negation and a well-chosen ε do the work, and ends with the laws that move a quantifier across a connective, and with prenex form, where every quantifier stands at the front.

Remark (Looking back). Section 3.2 proved two statements equivalent by two derivations, one each way. Every law of this section is such a pair. What is new is that a law may now be used on a part of a formula buried under quantifiers, and the subsection on negating a definition says why that is safe.

6.2.1 Not every, and some not

Section 5.1 called the two quantifiers mirror images: a single counterexample refutes a "for every", and a single witness confirms a "there exists". The mirror suggests four sequents. "Not every x satisfies φ" should say the same as "some x fails φ", and "no x satisfies φ" the same as "every x fails φ". Three of the four directions come straight out of the search method. The fourth does not, and the reason it does not is worth a close look.

To negate a quantifier, flip it and move the $\neg$ inside.

In the derivations of this section, $\neg\varphi[x_0/x]$ means $(\neg\varphi)[x_0/x]$, which is $\neg(\varphi[x_0/x])$ by Definition 5.2.5; the two readings give the same formula. The fresh name x_0 is always a variable different from x that does not appear in φ (or in ψ, when there is one); so it does not appear in $\forall x\,\varphi$ or $\exists x\,\varphi$ either. The side conditions of $\forall$I and $\exists$E need this, and it also makes x_0 free for x in φ (Section 5.2).

Proposition 6.2.1. Quantifier Negation

Let φ be a formula and x a variable.

(a) $\neg\forall x\,\varphi \vdash \exists x\,\neg\varphi$, by a derivation that uses RAA;
(b) $\exists x\,\neg\varphi \vdash \neg\forall x\,\varphi$, by a derivation that does not;
(c) $\neg\exists x\,\varphi \dashv\vdash \forall x\,\neg\varphi$, both directions by derivations that do not use RAA.

Try it first (ten minutes): run the search method of Section 3.1, with the quantifier steps of Sections 5.3–5.6, on (a). Which rule should give the last line, and what would that rule need from the premise?

Idea. Start with (a), and with the obstacle. The goal $\exists x\,\neg\varphi$ is an $\exists$, so the search postpones it until an available line shows an instance $\neg\varphi[t/x]$. None ever does. The premise $\neg\forall x\,\varphi$ is a negation, and a negation is used only by clashing it with what it negates, $\forall x\,\varphi$. It says that a counterexample exists and names none, so $\exists$I has nothing to act on. We are stuck, and the last step of the method applies: assume $\neg\exists x\,\neg\varphi$, "no x fails φ", and hunt for a clash. The clash will be with the premise, so the new goal is $\forall x\,\varphi$. Open a flagged box for it, with the goal $\varphi[x_0/x]$. That goal is again stuck (we know nothing about φ), so a second RAA assumes $\neg\varphi[x_0/x]$; and now an instance of $\neg\varphi$ is on a line, so $\exists$I produces $\exists x\,\neg\varphi$, which clashes with the outer assumption.

For (b) and (c) the method runs without getting stuck. A negated goal opens a $\neg$I box (Section 2.2); an $\exists$ premise opens an $\exists$E box early (Section 5.6); a $\forall$ goal opens a flagged

box (Section 5.4). The clash comes in one of two ways. In (b), and in (c) right to left, the instance assumed in the ∃E box clashes with the matching instance that ∀E takes from the ∀ line. In (c) left to right, ∃I turns the instance into $\exists x\,\varphi$, which meets the premise $\neg\exists x\,\varphi$.

Proof. For (a), the derivation:

1	$\neg\forall x\,\varphi$	
2	$\quad\neg\exists x\,\neg\varphi$	asm, for RAA
3	$\quad x_0 \quad \neg\varphi[x_0/x]$	asm, for RAA
4	$\quad\quad\exists x\,\neg\varphi$	∃I 3
5	$\quad\quad\bot$	¬E 4, 2
6	$\quad\varphi[x_0/x]$	RAA 3–5
7	$\quad\forall x\,\varphi$	∀I 3–6
8	$\quad\bot$	¬E 7, 1
9	$\exists x\,\neg\varphi$	RAA 2–8

Line 4 is ∃I at the term x_0, which is free for x in $\neg\varphi$. At line 7 the side condition of ∀I holds: x_0 does not appear in the premise, in the assumption $\neg\exists x\,\neg\varphi$ still open at line 7, or in $\forall x\,\varphi$, since it is not x and does not appear in φ; and line 6 is the last line of the flagged box.

For (b):

1	$\exists x\,\neg\varphi$	
2	$\quad\forall x\,\varphi$	asm, for ¬I
3	$\quad x_0 \quad \neg\varphi[x_0/x]$	asm, for ∃E
4	$\quad\quad\varphi[x_0/x]$	∀E 2
5	$\quad\quad\bot$	¬E 4, 3
6	$\quad\bot$	∃E 1, 3–5
7	$\neg\forall x\,\varphi$	¬I 2–6

At line 6, x_0 does not appear in $\bot$, in $\exists x\,\neg\varphi$, in the premise, or in the assumption $\forall x\,\varphi$ still open there.

For (c), left to right and right to left:

1	$\neg\exists x\,\varphi$	
2	x_0 | $\varphi[x_0/x]$	asm, for $\neg$I
3	$\exists x\,\varphi$	$\exists$I 2
4	$\bot$	$\neg$E 3, 1
5	$\neg\varphi[x_0/x]$	$\neg$I 2–4
6	$\forall x\,\neg\varphi$	$\forall$I 2–5

1	$\forall x\,\neg\varphi$	
2	$\exists x\,\varphi$	asm, for $\neg$I
3	x_0 | $\varphi[x_0/x]$	asm, for $\exists$E
4	$\neg\varphi[x_0/x]$	$\forall$E 1
5	$\bot$	$\neg$E 3, 4
6	$\bot$	$\exists$E 2, 3–5
7	$\neg\exists x\,\varphi$	$\neg$I 2–6

In the first, line 5 is the instance of $\neg\varphi$ at x_0, as $\forall$I requires, and x_0 appears neither in the premise nor in $\forall x\,\neg\varphi$; no assumption is open at line 6. In the second, the side condition of $\exists$E at line 6 is checked as in (b). No line of the derivations of (b) and (c) is justified by RAA, and the derivation of (a) uses it twice, at lines 6 and 9. This proves the proposition.

■

So "not every" and "some not" are interderivable, and so are "none" and "every not". The direction that cost two RAA boxes is the one that **produces** an existential statement from a negative one. That is not an accident of our derivation: no derivation of (a) avoids RAA. This is a claim about all derivations, which no search can establish, and we state it without proof (Chapter 4's optional Theorem 4.7.4 proves the propositional analogue for $\neg\neg P \vdash P$).

Quick check. From the line $\neg\forall x\,P(x)$, may a derivation write $\neg P(c)$ for a constant c? If not, what may it write, and what does a derivation do next to use it?

Solution. No. $\neg\forall x\,P(x)$ says that **some** element fails P, not that the element named c does: over $\mathbb{Z}$, with P read as "is even" and c as 0, $\neg\forall x\,P(x)$ is true and $\neg P(0)$ is false. By (a) the derivation may write $\exists x\,\neg P(x)$. To use that line it opens an $\exists$E box with a **fresh** witness x_0 and the assumption $\neg P(x_0)$ (Section 5.6), and nothing about c follows.

The laws as one line. From now on a derivation may use any law of this section as one line, citing it, in the way Proposition 6.1.1 allows a bounded rule: the line abbreviates the derivation printed here, with φ (and ψ) replaced by the formulas at hand and the fresh name chosen so that it appears nowhere in the larger derivation. Then every side condition inside the pasted block still holds, since the block's premises become lines of the larger derivation and the name is new to all of them. In a derivation the justification is a short tag: QN (a) for Proposition 6.2.1 (a), NIm for Proposition 6.2.2 below, and VI for Proposition 2.1.2.

6.2.2 The negation of an implication

Every definition with a bounded $\forall$ contains an implication, $\forall x\,(x \in A \to \varphi)$, so negating it means negating an implication. The answer is the one Section 5.1 found for the wrong translation of a bounded $\exists$: an implication fails exactly when its hypothesis holds and its conclusion does not. Chapter 3 met this law only as an exercise (Exercise (B1: Three searches) (b)), and a proof in the text may not rest on an exercise, so we prove it here.

Proposition 6.2.2. Negated Implication

For all formulas φ and ψ,

$$\neg(\varphi \to \psi) \dashv\vdash \varphi \wedge \neg\psi.$$

The direction $\vdash$ is derived with RAA, and the direction $\dashv$ without it.

Idea. For $\vdash$ the goal is a conjunction, so $\wedge$I asks for φ and $\neg\psi$ separately. The premise is a negation, usable only to clash with $\varphi \to \psi$, so in each part we manufacture $\varphi \to \psi$. For $\neg\psi$: open a $\neg$I box on ψ; then $\varphi \to \psi$ follows by $\to$I, whose box repeats ψ. For φ we are stuck, since φ has no known main connective; so RAA assumes $\neg\varphi$, and $\neg\varphi$ gives $\varphi \to \psi$ by Proposition 2.1.2. For $\dashv$, the goal is a negation: assume $\varphi \to \psi$ and clash ψ with $\neg\psi$.

Proof. The two derivations:

1	$\neg(\varphi \to \psi)$	
2	$\quad\neg\varphi$	asm, for RAA
3	$\quad\varphi \to \psi$	VI 2
4	$\quad\bot$	¬E 3, 1
5	φ	RAA 2–4
6	$\quad\psi$	asm, for ¬I
7	$\quad\quad\varphi$	asm, for →I
8	$\quad\quad\psi$	R 6
9	$\quad\varphi \to \psi$	→I 7–8
10	$\quad\bot$	¬E 9, 1
11	$\neg\psi$	¬I 6–10
12	$\varphi \wedge \neg\psi$	∧I 5, 11

1	$\varphi \wedge \neg\psi$	
2	$\quad\varphi \to \psi$	asm, for ¬I
3	$\quad\varphi$	∧E 1
4	$\quad\psi$	→E 2, 3
5	$\quad\neg\psi$	∧E 1
6	$\quad\bot$	¬E 4, 5
7	$\neg(\varphi \to \psi)$	¬I 2–6

Line 3 of the first derivation uses Proposition 2.1.2 as a derived rule (Definition 2.3.1), and line 8 repeats line 6 into the box, since →I needs ψ as a line of its own box. The only RAA is at line 5.

■

The derivations use only the rules of Chapters 1 and 2, so the law holds for statements as well as formulas, and it may be used as a derived rule in the sense of Definition 2.3.1. Over the real numbers: "if $x > 2$ then $x^2 > 4$" fails for a given x exactly when $x > 2$ and $x^2 \leq 4$.

6.2.3 Bounded quantifiers keep their bounds

Analysis negates bounded quantifiers: $\neg\forall\varepsilon > 0$, $\neg\exists N \in \mathbb{N}$. The two laws so far combine into laws for them, and the combination has one surprise in it, which is no surprise once the bound is unfolded: **the bound stays as it is**.

Negating a bounded quantifier flips the quantifier and keeps the bound.

Proposition 6.2.3. Negating Bounded Quantifiers

Let A be a fixed set, so that $\in A$ is a one-place predicate symbol, let x be a variable and φ a formula. Then

(a) $\neg\forall x \in A\, \varphi \dashv\vdash \exists x \in A\, \neg\varphi$, where $\vdash$ is derived with RAA and $\dashv$ without;
(b) $\neg\exists x \in A\, \varphi \dashv\vdash \forall x \in A\, \neg\varphi$, both directions without RAA.

The same holds with $x \in A$ replaced by $x > a$, or by $x \geq a$, $x < a$, $x \leq a$, for a term a in which x does not occur.

So $\neg\forall \varepsilon > 0\, \varphi$ becomes $\exists \varepsilon > 0\, \neg\varphi$, and never $\exists \varepsilon \leq 0\, \neg\varphi$.

Idea. Unfold the bounds (Definition 5.1.1): (a) is $\neg\forall x\,(x \in A \to \varphi) \dashv\vdash \exists x\,(x \in A \wedge \neg\varphi)$. Push the $\neg$ through the quantifier by Proposition 6.2.1, and then through the implication by Proposition 6.2.2; the second step must happen under the $\exists$, so it happens inside an $\exists$E box. The negation lands on φ, and the bound $x \in A$ comes out of the implication unnegated. For (b), the derivations of Proposition 6.2.1 (c) go through with the bounded rules of Proposition 6.1.1 in place of the unbounded ones.

Proof. For (a), $\vdash$ and $\dashv$, with the bounds unfolded:

1		$\neg\forall x\,(x \in A \to \varphi)$	
2		$\exists x\,\neg(x \in A \to \varphi)$	QN (a) 1
3	x_0	$\neg(x_0 \in A \to \varphi[x_0/x])$	asm, for $\exists$E
4		$x_0 \in A \wedge \neg\varphi[x_0/x]$	NIm 3
5		$\exists x\,(x \in A \wedge \neg\varphi)$	$\exists$I 4
6		$\exists x\,(x \in A \wedge \neg\varphi)$	$\exists$E 2, 3–5

1		$\exists x\,(x \in A \wedge \neg\varphi)$	
2	x_0	$x_0 \in A \wedge \neg\varphi[x_0/x]$	asm, for $\exists$E
3		$\neg(x_0 \in A \to \varphi[x_0/x])$	NIm 2
4		$\exists x\,\neg(x \in A \to \varphi)$	$\exists$I 3
5		$\exists x\,\neg(x \in A \to \varphi)$	$\exists$E 1, 2–4
6		$\neg\forall x\,(x \in A \to \varphi)$	QN (b) 5

Line 3 on the left is the instance at x_0 of $\neg(x \in A \to \varphi)$, since A is not a variable; line 4 is Proposition 6.2.2 with $x_0 \in A$ and $\varphi[x_0/x]$ for φ and ψ. At line 6, x_0 does not appear in the conclusion, in line 2 or in the premise, and no assumption is open. The right-hand derivation is checked the same way. The only use of RAA is inside the two cited laws on the left: in QN (a) and in the direction $\vdash$ of NIm.

For (b), with the bounded rules:

1	$\neg\exists x \in A\ \varphi$	
2	x_0 | $x_0 \in A$	asm, for $\forall$I
3	$\varphi[x_0/x]$	asm, for $\neg$I
4	$\exists x \in A\ \varphi$	$\exists$I 2, 3
5	$\bot$	$\neg$E 4, 1
6	$\neg\varphi[x_0/x]$	$\neg$I 3–5
7	$\forall x \in A\ \neg\varphi$	$\forall$I 2–6

1	$\forall x \in A\ \neg\varphi$	
2	$\exists x \in A\ \varphi$	asm, for $\neg$I
3	x_0 | $x_0 \in A$	asm, for $\exists$E
4	$\varphi[x_0/x]$	asm, for $\exists$E
5	$\neg\varphi[x_0/x]$	$\forall$E 1, 3
6	$\bot$	$\neg$E 4, 5
7	$\bot$	$\exists$E 2, 3–6
8	$\neg\exists x \in A\ \varphi$	$\neg$I 2–7

At line 7 on the left, the box's own assumption $x_0 \in A$ has closed, and x_0 is in neither the premise nor the conclusion; at line 7 on the right, x_0 is not in $\bot$, in line 2, in the premise, or in the open assumption on line 2. Neither derivation uses RAA.

For the bounds $x > a$ and the others, the same four derivations work with $x > a$ and $x_0 > a$ in place of $x \in A$ and $x_0 \in A$, choosing x_0 not to appear in a; since x does not occur in a, $(x > a)[x_0/x]$ is $x_0 > a$. This proves the proposition.

■

Warning. **Negate the quantifier, not the bound.** The negation of "$a_n \to L$" begins $\exists\varepsilon > 0$, not $\exists\varepsilon \le 0$. The wrong form, $\exists\varepsilon \le 0\ \forall N \in \mathbb{N}\ \exists n \in \mathbb{N}\ (n \ge N \wedge |a_n - L| \ge \varepsilon)$, is true for **every** sequence: $\varepsilon = 0$ works, since $|a_n - L| \ge 0$ always. So it would "prove" that the constant sequence $a_n = 0$ does not converge to 0.

6.2.4 Negating a definition, mechanically

With the laws in hand, negating a definition is a procedure.

Algorithm. **Negating a formula.**

1. **Rewrite** $\leftrightarrow$ as two implications, by Proposition 2.4.1.
2. **Push** $\neg$ **inward**, one symbol at a time, from the outside, citing a law at each step:
   - $\neg\forall x$ becomes $\exists x\,\neg$, and $\neg\exists x$ becomes $\forall x\,\neg$ (Proposition 6.2.1); a bounded quantifier keeps its bound (Proposition 6.2.3);
   - $\neg(\alpha \to \beta)$ becomes $\alpha \wedge \neg\beta$ (Proposition 6.2.2);
   - $\neg(\alpha \wedge \beta)$ becomes $\neg\alpha \vee \neg\beta$, and $\neg(\alpha \vee \beta)$ becomes $\neg\alpha \wedge \neg\beta$ (De Morgan, Proposition 2.5.5);
   - $\neg\neg\alpha$ becomes α (Proposition 2.5.1).
3. **Stop at atomic formulas.** In an ordered domain, replace $\neg(s < t)$ by $s \ge t$ and $\neg(s \le t)$ by $s > t$; these are facts about the order, not laws of logic.

Each step of 1 and 2 replaces a part of the formula by a part provably equivalent to it. Write that chain with $\dashv\vdash$ between the stages, not $\equiv$. Step 3 is not a law of logic: make its replacements after the chain, citing the order fact.

Why may a step act on a part buried inside the formula? Because a part may be replaced by a part provably equivalent to it, even under quantifiers that bind its variables, as long as the equivalence is derivable for every choice of those variables, that is, its closure is derivable.

Proposition 6.2.4. Replacement of Provable Equivalents

Let φ and ψ be formulas, and let $z_1, \dots, z_k$ be the variables free in φ or in ψ. Suppose that $\vdash \forall z_1 \cdots \forall z_k\,(\varphi \leftrightarrow \psi)$. If χ' is obtained from a formula χ by replacing one occurrence of φ in χ by ψ, then $\vdash \forall w_1 \cdots \forall w_m\,(\chi \leftrightarrow \chi')$, where $w_1, \dots, w_m$ are the variables free in χ or in χ'. In particular, $\chi \dashv\vdash \chi'$.

In words: swap a part for a provably equivalent part anywhere in a formula, provided the equivalence is derivable for every choice of its free variables, that is, its closure is derivable; every law of this section is.

The proof is an induction on the formula χ; its connective steps are short derivations with the rules of Chapters 1 and 2, its quantifier steps use a flagged box or an $\exists$E box with a fresh name, and it does not use truth tables or completeness. It is in the fold below, followed by a second fold that checks the hypothesis for the laws of this section.

Second reading · *The proof of the replacement rule*

Proof. For a formula θ, write $\overline{\forall}\theta$ for its **closure**: θ preceded by one $\forall$ for each variable free in θ. The order of these $\forall$'s does not matter, because the instance of $\forall z_1 \cdots \forall z_k\,\theta$ at fresh flags comes by $\forall$E in that order, and k flagged boxes closed by $\forall$I in any other order give the $\forall$'s in that other order (this is the derivation of Example (Flagged boxes at work) (c), with k variables), so "the closure" means the closure in any one order. The hypothesis is $\vdash \overline{\forall}(\varphi \leftrightarrow \psi)$, and the claim is $\vdash \overline{\forall}(\chi \leftrightarrow \chi')$. A derivable formula may stand as a line of any derivation: paste its derivation in, with its flags renamed to names new to the whole derivation, as "The laws as one line" allows.

We use strong induction (Theorem 4.2.2) on the number of connectives and quantifier symbols of χ: assume the claim for every formula with fewer such symbols, and prove it for χ.

Case 0: the occurrence is χ itself. Then χ is φ, χ' is ψ, and the claim is the hypothesis.

Otherwise the occurrence lies in an immediate part of χ, which has fewer symbols; an atomic formula has no part other than itself.

Case 1: $\chi = \neg\chi_0$, with the occurrence in χ_0, so $\chi' = \neg\chi_0'$. The induction hypothesis gives $\vdash \overline{\forall}(\chi_0 \leftrightarrow \chi_0')$. Open one flagged box for each variable of the closure of $\neg\chi_0 \leftrightarrow \neg\chi_0'$, nested in the closure's order, with fresh flags. Inside, $\forall$E of the induction hypothesis at the flags, each variable at its own flag in the closure's order, gives the instance $\alpha \leftrightarrow \alpha'$ of $\chi_0 \leftrightarrow \chi_0'$ at the flags; every variable of that closure is free in χ or χ', so it has a flag. The rules of Sections 2.2 and 2.4 give $\neg\alpha \leftrightarrow \neg\alpha'$: each direction is a $\neg$I box (Definition 2.2.1) with $\leftrightarrow$E inside, and $\leftrightarrow$I (Definition 2.4.1) joins them. The flagged boxes then close by $\forall$I, the innermost first. Each side condition holds: the flag is fresh, and no assumption is open at the closing line.

Case 2: $\chi = \chi_1 \circ \chi_2$, where $\circ$ is $\wedge$, $\vee$, $\rightarrow$ or $\leftrightarrow$, and the occurrence is in χ_1 (in χ_2, exchange the parts). As in case 1, with the propositional step $\alpha \leftrightarrow \alpha' \vdash (\alpha \circ \gamma) \leftrightarrow (\alpha' \circ \gamma)$, where γ

is the instance of χ_2 at the flags. For each $\circ$ it is a short derivation with that connective's rules of Chapters 1 and 2 and $\leftrightarrow$ E on $\alpha \leftrightarrow \alpha'$. A variable free in χ but not in χ_1 is flagged and never instantiated, which is harmless: the instance of the induction hypothesis is taken at the flags of its own variables only.

Case 3: $\chi = \forall x\, \chi_0$ or $\exists x\, \chi_0$, with the occurrence in χ_0, so χ' is $\forall x\, \chi_0'$ or $\exists x\, \chi_0'$. The induction hypothesis gives $\vdash \overline{\forall}(\chi_0 \leftrightarrow \chi_0')$, whose closure may include x. Let $\bar{w}$ be the variables free in χ or χ', so x is not among them, flag them with fresh names $\bar{w}_0$, and write θ_0 for $\theta[\bar{w}_0/\bar{w}]$. Substitution passes $\forall x$ and $\exists x$, because x is not among $\bar{w}$; and all the flags are fresh, so the simultaneous and the step-by-step substitutions agree.

- For $\forall$, left to right: assume $\forall x\, (\chi_0)_0$ and open a flagged box with a fresh x_1. Use $\forall$E at x_1. Use $\forall$E of the induction hypothesis at $\bar{w}_0$ and x_1, each variable at its own flag in the closure's order. Then $\leftrightarrow$ E gives $(\chi_0')_0[x_1/x]$, and $\forall$I gives $\forall x\, (\chi_0')_0$. The flag x_1 is in no premise, open assumption or conclusion.
- Right to left is the same, and $\leftrightarrow$ I joins the two directions.
- For $\exists$: an $\exists$E box with a fresh witness x_1, $\leftrightarrow$ E inside, closed after $\exists$I.

The flagged boxes of $\bar{w}_0$ then close by $\forall$I, as in case 1.

The "in particular". From $\vdash \overline{\forall}(\chi \leftrightarrow \chi')$, $\forall$E at each w_i itself (a variable is free for itself) gives $\chi \leftrightarrow \chi'$, and $\leftrightarrow$ E gives each direction. No truth table and no completeness theorem was used.

■

Second reading · *Why the laws of this section meet the hypothesis*

A law of this section is proved for every formula that meets its condition, so it holds for the instances at fresh names $z_{1,0}, \ldots, z_{k,0}$, and $\leftrightarrow$ I (Definition 2.4.1) joins the two directions into $\alpha \leftrightarrow \beta$ at the fresh names, and k flagged boxes give its closure, as Proposition 6.2.4 asks. Take one variable x at a time, with sides α and β of the law, and a fresh x_0: a variable other than x that appears in neither α nor β. The derivation of $\alpha[x_0/x] \vdash \beta[x_0/x]$ is available because each law used holds for **every** formula that meets its condition, in particular for the instances at x_0. The laws this book applies to parts have conditions of two kinds, and both survive, for the same reason; the passage laws at the end of this section have one kind, checked in the fold after them, and the bounded negation laws the other. A law of Proposition 6.2.3 with a bound such as $z > a$ asks that its bound variable z not occur in the term a; since z appears in β, x_0 is not z (as for the passage laws, in the fold after them). If z is x, substitution stops at the quantifier on x and the pair is unchanged; otherwise it passes the quantifier on z, and $a[x_0/x]$ contains no variable that a lacks except x_0, so z does not occur in it. So the instance is again an instance of the same law, with its condition met.

Renaming a bound variable is a replacement of this kind: Section 5.4 derived $\forall y\, \varphi \dashv\vdash \forall z\, \varphi[z/y]$ for every formula φ in which z does not appear (the fold after Example (Flagged boxes at work)), and Section 5.6 did the same for $\exists$, so the convention of Section 5.2, that a renaming may be made anywhere inside a formula, is now a provable equivalence. Its closure is derivable, as Proposition 6.2.4 asks: for a variable x free in it and a fresh x_0 (as in the fold above), x is

neither y nor z, and x_0 is not z, because z appears in $\forall z\,\varphi[z/y]$, so z does not appear in $\varphi[x_0/x]$, and the instance is the renaming of $\forall y\,\varphi[x_0/x]$ to $\forall z\,\varphi[x_0/x][z/y]$. Two formulas that differ only by renaming bound variables are joined by a chain of single renamings, each to a letter absent from the formula at that step (for $\forall x\,\forall y\,R(x,y)$ and $\forall y\,\forall x\,R(y,x)$: x to u, then y to x, then u to y), each applied to a part by Proposition 6.2.4 and chained by Lemma 2.3.3.

Example. Three negations from analysis

Negate each statement into positive form: no $\neg$ in front of a quantifier, a connective or an inequality. Work over $\mathbb{R}$.

(a) $a_n \to L$: $\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to |a_n - L| < \varepsilon)$.

(b) s is the supremum of A, in the two-clause form "s is an upper bound of A, and for every $\varepsilon > 0$ some element of A exceeds $s - \varepsilon$":

$$\forall x \in A\ (x \leq s)\ \wedge\ \forall \varepsilon > 0\ \exists x \in A\ (s - \varepsilon < x).$$

(c) f is uniformly continuous on A: $\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A\ \forall y \in A\ (|x - y| < \delta \to |f(x) - f(y)| < \varepsilon)$.

Solution. (a) Three quantifiers, then an implication, then an inequality:

$$\begin{aligned}
&\neg\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to |a_n - L| < \varepsilon)\\
\dashv\vdash\ &\exists \varepsilon > 0\ \neg\exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to |a_n - L| < \varepsilon)\\
\dashv\vdash\ &\exists \varepsilon > 0\ \forall N \in \mathbb{N}\ \neg\forall n \in \mathbb{N}\ (n \geq N \to |a_n - L| < \varepsilon)\\
\dashv\vdash\ &\exists \varepsilon > 0\ \forall N \in \mathbb{N}\ \exists n \in \mathbb{N}\ \neg(n \geq N \to |a_n - L| < \varepsilon)\\
\dashv\vdash\ &\exists \varepsilon > 0\ \forall N \in \mathbb{N}\ \exists n \in \mathbb{N}\ (n \geq N \wedge \neg(|a_n - L| < \varepsilon)).
\end{aligned}$$

The first three steps are Proposition 6.2.3, the fourth is Proposition 6.2.2. Last, the order fact turns $\neg(|a_n - L| < \varepsilon)$ into $|a_n - L| \geq \varepsilon$:

$$\exists \varepsilon > 0\ \forall N \in \mathbb{N}\ \exists n \in \mathbb{N}\ (n \geq N \wedge |a_n - L| \geq \varepsilon).$$

In words: some $\varepsilon > 0$ is such that, however far out we start, some later term is at least ε away from L. Every bound stayed, and the order of the quantifiers stayed: ε, then N, then n.

(b) The outer symbol is $\wedge$. By De Morgan the negation becomes a disjunction, and each side is negated on its own:

$$\begin{aligned}
&\neg\big(\forall x \in A\ (x \leq s)\ \wedge\ \forall \varepsilon > 0\ \exists x \in A\ (s - \varepsilon < x)\big)\\
\dashv\vdash\ &\neg\forall x \in A\ (x \leq s)\ \vee\ \neg\forall \varepsilon > 0\ \exists x \in A\ (s - \varepsilon < x)\\
\dashv\vdash\ &\exists x \in A\ \neg(x \leq s)\ \vee\ \exists \varepsilon > 0\ \forall x \in A\ \neg(s - \varepsilon < x).
\end{aligned}$$

The first step is Proposition 2.5.5, (d) for $\vdash$ and (c) for $\dashv$; the second is Proposition 6.2.3, (a) on the left and (a) then (b) on the right. Last, the order facts turn $\neg(x \leq s)$ into $s < x$ and $\neg(s - \varepsilon < x)$ into $x \leq s - \varepsilon$:

$$\exists x \in A\ (s < x)\ \vee\ \exists \varepsilon > 0\ \forall x \in A\ (x \leq s - \varepsilon).$$

In words: s fails to be the supremum when some element of A lies above it (it is not an upper bound), or when some $s - \varepsilon$ with $\varepsilon > 0$ is still an upper bound (it is not the least).

(c) Four quantifiers and an implication:

$$\begin{aligned} &\neg\forall\varepsilon > 0\ \exists\delta > 0\ \forall x \in A\ \forall y \in A\ \big(|x - y| < \delta \to |f(x) - f(y)| < \varepsilon\big) \\ \dashv\vdash\ &\exists\varepsilon > 0\ \forall\delta > 0\ \exists x \in A\ \exists y \in A\ \neg\big(|x - y| < \delta \to |f(x) - f(y)| < \varepsilon\big) \\ \dashv\vdash\ &\exists\varepsilon > 0\ \forall\delta > 0\ \exists x \in A\ \exists y \in A\ \big(|x - y| < \delta \wedge \neg(|f(x) - f(y)| < \varepsilon)\big), \end{aligned}$$

where the first step is four uses of Proposition 6.2.3, and the second is Proposition 6.2.2. Last, the order fact turns $\neg(|f(x) - f(y)| < \varepsilon)$ into $|f(x) - f(y)| \geq \varepsilon$:

$$\exists\varepsilon > 0\ \forall\delta > 0\ \exists x \in A\ \exists y \in A\ \big(|x - y| < \delta \wedge |f(x) - f(y)| \geq \varepsilon\big).$$

In words: one $\varepsilon > 0$ defeats every δ: however small $\delta > 0$ is, two points of A closer than δ have values at least ε apart. To prove that a function is **not** uniformly continuous, this is the statement to prove, and it says exactly what to exhibit: first ε; then, for an arbitrary δ, the two points x and y, which may depend on δ. Section 6.4 does this for x^2 on $\mathbb{R}$.

The same procedure finishes the rewriting promised in Section 6.1 for "there is no largest prime". By Proposition 6.2.1 (c),

$$\begin{aligned} &\neg\exists m\ \big(\mathrm{Prime}(m) \wedge \forall n\ (\mathrm{Prime}(n) \to n \leq m)\big) \\ \dashv\vdash\ &\forall m\ \neg\big(\mathrm{Prime}(m) \wedge \forall n\ (\mathrm{Prime}(n) \to n \leq m)\big). \end{aligned}$$

A negated conjunction $\neg(\alpha \wedge \beta)$ is interderivable with $\alpha \to \neg\beta$: De Morgan's law (Proposition 2.5.5 (d)) and Proposition 2.5.6 (a) give one direction, and (b) of the latter with Proposition 2.5.5 (c) the other. Then Proposition 6.2.1 (a) and (b) turn $\neg\forall n\,(\mathrm{Prime}(n) \to n \leq m)$ into $\exists n\,\neg(\mathrm{Prime}(n) \to n \leq m)$, and Proposition 6.2.2, used under $\exists n$, pushes the last $\neg$ in; both steps act on a part, by Proposition 6.2.4:

$$\forall m\ \Big(\mathrm{Prime}(m) \to \exists n\ \big(\mathrm{Prime}(n) \wedge \neg(n \leq m)\big)\Big).$$

The quantifier $\forall n\,(\mathrm{Prime}(n) \to \dots)$ works like a bounded one, with the bound $\mathrm{Prime}(n)$. That bound is not one that Proposition 6.2.3 covers, but the same two steps work for it, and the bound stays unnegated.

Quick check. Negate "every non-empty set of reals that is bounded above has a supremum", writing $E(A)$ for "A is non-empty", $B(A)$ for "A is bounded above" and $S(A)$ for "A has a supremum", with the variable A ranging over sets of reals.

Solution. The statement is $\forall A\,\big((E(A) \wedge B(A)) \to S(A)\big)$. By Proposition 6.2.1 (a) and (b) and then Proposition 6.2.2, its negation is

$$\exists A\ \big((E(A) \wedge B(A)) \wedge \neg S(A)\big)\ :$$

there is a non-empty set of reals, bounded above, with no supremum. The two conditions are not negated: they stay as conditions on the counterexample, just as a bound stays with its quantifier.

Quick check. Negate "f is bounded on A", $\exists M\ \forall x \in A\ (|f(x)| \le M)$, into positive form, and say it in words.

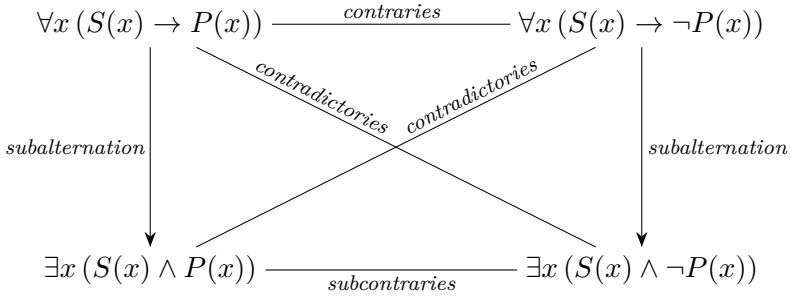
Solution. $\forall M\ \exists x \in A\ (|f(x)| > M)$: every proposed bound is beaten somewhere on A. The quantifiers flip, the bound $x \in A$ stays, and the order fact turns $\neg(|f(x)| \le M)$ into $|f(x)| > M$.

Warning. **The negation of an implication is not an implication.** "If $x > 2$, then $x^2 > 4$" is not negated by "if $x > 2$, then $x^2 \le 4$", nor by "if $x \le 2$, then $x^2 \le 4$". For $x = 0$ the original is true (its hypothesis fails), and so is the second candidate, so the candidate cannot be its negation. The negation is $x > 2 \wedge x^2 \le 4$ (Proposition 6.2.2), and it is false for every real x.

6.2.5 Four forms and their square

Four sentence shapes come up so often that they have names. Fix two one-place predicates S and P.

- **A**, "every S is P": $\forall x\,(S(x) \to P(x))$;
- **E**, "no S is P": $\forall x\,(S(x) \to \neg P(x))$;
- **I**, "some S is P": $\exists x\,(S(x) \wedge P(x))$;
- **O**, "some S is not P": $\exists x\,(S(x) \wedge \neg P(x))$.



The diagonals always hold: A and O are each the negation of the other, and so are E and I. The other three labels hold only when something is an S.

The diagonals, by the laws. A–O is the two steps of the proof of Proposition 6.2.3 (a), QN and then NIm under $\exists$, which work for the bound $S(x)$ as for Prime(n) above. E–I is the same, followed by Proposition 2.5.1 under $\exists x$ (by Proposition 6.2.4).

Contrary against contradictory. A and E are **contrary**: never both true when something is an S, but possibly both false. The negation of "every element of A is positive" is the O form "some element of A is not positive", not the E form "every element of A is non-positive".

Example. The square for multiples of 4 and squares

Over $\mathbb{N}$, let S be "is a multiple of 4" and P "is a perfect square". Find the truth value of each corner. Then do the same with S read as "is an odd multiple of 4".

Solution. A, "every multiple of 4 is a perfect square", is false at 8. E, "no multiple of 4 is a perfect square", is false at $4 = 2^2$. I is true at 4, and O at 8. Each diagonal joins a true corner to a false one.

No number is an odd multiple of 4. Now A and E are both true, vacuously, and I and O both false. Contraries both true, subcontraries both false: this is why the outer labels need something to be an S (existential import, Section 5.1).

Quick check. Which corner is the negation of "no multiple of 4 is a perfect square", and is that corner true?

Solution. I: "some multiple of 4 is a perfect square". It is true, since $16 = 4 \cdot 4 = 4^2$. The negation of E is I, by the diagonal.

6.2.6 How the laws look in a mathematical proof

Chapter 2 proved, in Example (A number below every positive number), that a number $x \geq 0$ below every positive number is 0. Its key step chose $\varepsilon := x/2$, and the reading after it called "the hypothesis applies to it" a step about "every", used as ordinary reasoning until Chapter 5. The rule is ∀E, in its bounded form (Proposition 6.1.1), at a term built from the data. Here is the same move once more, the way it appears in analysis.

Two real numbers that are closer than every positive number are equal.

Proposition 6.2.5. Arbitrarily Close Numbers Are Equal

Let a and b be real numbers. If $|a - b| < \varepsilon$ for every $\varepsilon > 0$, then $a = b$.

Idea. The goal $a = b$ is atomic, so no introduction rule applies, and the hypothesis is a ∀ whose instances are all inequalities, never an equation. We are stuck, so suppose $a \neq b$. The hypothesis may be instantiated at any positive term, so choose the term that makes the numbers collide: $\varepsilon := |a - b|$, which is positive exactly because $a \neq b$.

Proof. Suppose that $|a - b| < \varepsilon$ for every $\varepsilon > 0$, and suppose, for a contradiction, that $a \neq b$ [an RAA box]. Then $a - b \neq 0$, so $|a - b| > 0$, by the school fact that $|u| > 0$ for every real $u \neq 0$. So the hypothesis applies to the term $|a - b|$ [bounded ∀E, citing the line $|a - b| > 0$] and gives $|a - b| < |a - b|$. No real number is less than itself, so this is a contradiction [¬E]. Hence $a = b$ [RAA].

■

The last step is a genuine RAA: the box assumed $\neg(a = b)$, written $a \neq b$, and concluded $a = b$; ¬I would only have given $\neg\neg(a = b)$. The hypothesis cannot be weakened to "for **some** $\varepsilon > 0$": $a = 0$, $b = \frac{1}{2}$, $\varepsilon = 1$ satisfy $|a - b| < \varepsilon$ and $a \neq b$. The facts about $|\cdot|$ used here are school algebra.

Read the result through the negation laws. Its contrapositive says: if $a \neq b$, then $\neg\forall\varepsilon > 0\,(|a-b| < \varepsilon)$, which by Proposition 6.2.3 is interderivable with $\exists\varepsilon > 0\,\neg(|a-b| < \varepsilon)$, that is, by the order fact, $\exists\varepsilon > 0\,(|a-b| \geq \varepsilon)$. The proof did not search for that ε: it named it, $|a-b|$, from the data. A proof by contradiction in analysis is usually this shape. Negate the goal; the negation hands over an object (here the positive number $|a-b|$); feed it to a hypothesis that holds for **every** object of its kind.

6.2.7 Moving a quantifier across a connective

Negation is not the only thing that needs moving. A formula such as $\exists x\,P(x) \to \psi$ has its quantifier buried in a hypothesis, and proofs often want it outside, or the reverse. Two groups of laws do this. The **distribution** laws move a quantifier across $\wedge$ or $\vee$ when both sides mention x; the **passage** laws move it across a connective when one side does not. The laws of this subsection are proved by the method's short runs; the derivations are in a fold, and Section 7.2 returns to two of the one-way laws.

Proposition 6.2.6. Distribution of Quantifiers

For all formulas φ, ψ and every variable x:

(a) $\forall x\,(\varphi \wedge \psi) \dashv\vdash \forall x\,\varphi \wedge \forall x\,\psi$;
(b) $\exists x\,(\varphi \vee \psi) \dashv\vdash \exists x\,\varphi \vee \exists x\,\psi$;
(c) $\forall x\,\varphi \vee \forall x\,\psi \vdash \forall x\,(\varphi \vee \psi)$;
(d) $\exists x\,(\varphi \wedge \psi) \vdash \exists x\,\varphi \wedge \exists x\,\psi$.

All four are derived without RAA.

The proposition collects laws that are mostly proved already. Its point is the comparison: $\forall$ goes across $\wedge$ and $\exists$ across $\vee$ in both directions, while $\forall$ across $\vee$ and $\exists$ across $\wedge$ go one way only.

Idea. Three of the four laws were derived in Sections 5.4 and 5.6 for $P(x)$ and $Q(x)$; only the fresh name's instances were used there, so formulas can take their place. For (c), the goal is a $\forall$ and the premise a disjunction: open the flagged box first, then split on the premise inside it.

The converses of (c) and (d) are false. Over $\mathbb{Z}$, every integer is even or odd, but neither "every integer is even" nor "every integer is odd" holds; some integer is even and some integer is odd, but none is both. These are counter-readings; Section 7.2 draws them as counter-models.

The passage laws need a condition: the quantified variable must not be free in the formula that the quantifier moves away from. Section 5.2 proved the fact that makes this work: if x is not free in ψ, then $\psi[t/x] = \psi$ for every term t. So an instance of $\psi \to \varphi$ at x_0 is $\psi \to \varphi[x_0/x]$, with ψ untouched.

Proposition 6.2.7. Passage Laws

Let φ and ψ be formulas and x a variable that is **not free in** ψ. Then

(a) $\forall x\,(\psi \to \varphi) \dashv\vdash \psi \to \forall x\,\varphi$;

(b) $\forall x\,(\varphi \to \psi) \dashv\vdash \exists x\,\varphi \to \psi$;
(c) $\exists x\,(\varphi \wedge \psi) \dashv\vdash \exists x\,\varphi \wedge \psi$;
(d) $\exists x\,(\varphi \to \psi) \dashv\vdash \forall x\,\varphi \to \psi$;
(e) $\exists x\,(\psi \to \varphi) \dashv\vdash \psi \to \exists x\,\varphi$.

Of the ten directions, two are derived here with RAA: $\dashv$ in (d), through LEM and the cited QN (a), and $\dashv$ in (e), through LEM. The other eight are derived without it.

Read (b) aloud: "everything with φ has ψ" says the same as "if something has φ, then ψ". In (b) and (d) the quantifier changes from $\forall$ to $\exists$ or back as it leaves the hypothesis of an implication. That is the game reading of Section 5.1 once more: a quantifier in a hypothesis changes sides.

Idea. Each intuitionistic direction is a short run of the search method. A goal $\to$ opens a $\to$I box; a goal $\forall$ a flagged box; a premise $\exists$ an $\exists$E box, early; a premise $\forall$ is instantiated at x_0, where the condition on ψ keeps ψ unchanged; and the $\exists$ goal waits for its witness x_0. The two classical directions get stuck. In (d), the premise $\forall x\,\varphi \to \psi$ is usable only once $\forall x\,\varphi$ is in hand, and it may not be; so decide it, by LEM. If $\forall x\,\varphi$ holds, then ψ holds, and $\varphi \to \psi$ holds at every x. If not, (a) of Proposition 6.2.1 supplies an x_0 with $\neg\varphi[x_0/x]$, and $\varphi[x_0/x] \to \psi$ is vacuously true. In (e), decide ψ: if ψ holds, the premise supplies a witness; if not, $\psi \to \varphi$ is vacuously true at every x.

Second reading · *The derivations of the distribution and passage laws*

Proof. For Proposition 6.2.6: (a) is Proposition 5.4.1, (d) is Proposition 5.6.1, and (b) is part (a) of Example (Derivations with existential elimination), the derivation of $\exists$ over $\vee$. Those derivations are written with $P(x)$ and $Q(x)$; they work word for word with φ and ψ in their place, because they use nothing about $P(x)$ and $Q(x)$ beyond the instances at the fresh name, which is chosen different from x and not appearing in φ or ψ.

For (c), the goal is a $\forall$: open a flagged box. Inside it the premise is a disjunction, so split on it; in each case one $\forall$E at x_0 and one $\vee$I reach the instance of $\varphi \vee \psi$:

1	$\forall x\,\varphi \vee \forall x\,\psi$		
2	x_0 | $\forall x\,\varphi$		asm, for $\vee$E
3	| $\varphi[x_0/x]$		$\forall$E 2
4	| $\varphi[x_0/x] \vee \psi[x_0/x]$		$\vee$I 3
5	| $\forall x\,\psi$		asm, for $\vee$E
6	| $\psi[x_0/x]$		$\forall$E 5
7	| $\varphi[x_0/x] \vee \psi[x_0/x]$		$\vee$I 6
8	$\varphi[x_0/x] \vee \psi[x_0/x]$		$\vee$E 1, 2–4, 5–7
9	$\forall x\,(\varphi \vee \psi)$		$\forall$I 2–8

Line 8 is the instance of $\varphi \vee \psi$ at x_0 and the last line of the flagged box; at line 9, x_0 is in neither

the premise nor the conclusion, and no assumption is open. This proves Proposition 6.2.6.

■

Proof. In every derivation below, x_0 is a variable other than x that does not appear in φ or ψ, and $\psi[x_0/x] = \psi$ because x is not free in ψ (Section 5.2). So the instance at x_0 of $\psi \to \varphi$ is $\psi \to \varphi[x_0/x]$, and likewise for the other combinations.

(a), $\vdash$ and $\dashv$:

1	$\forall x\,(\psi \to \varphi)$	
2	$\quad \psi$	asm, for $\to$I
3	$\quad x_0 \mid \psi \to \varphi[x_0/x]$	$\forall$E 1
4	$\quad\quad \varphi[x_0/x]$	$\to$E 3, 2
5	$\quad \forall x\,\varphi$	$\forall$I 3–4
6	$\psi \to \forall x\,\varphi$	$\to$I 2–5

1	$\psi \to \forall x\,\varphi$	
2	$x_0 \mid \psi$	asm, for $\to$I
3	$\quad \forall x\,\varphi$	$\to$E 1, 2
4	$\quad \varphi[x_0/x]$	$\forall$E 3
5	$\psi \to \varphi[x_0/x]$	$\to$I 2–4
6	$\forall x\,(\psi \to \varphi)$	$\forall$I 2–5

At line 5 on the left the assumption ψ is open, and x_0 does not appear in it: this is where the fresh name must avoid ψ.

(b), $\vdash$ and $\dashv$:

1	$\forall x\,(\varphi \to \psi)$	
2	$\quad \exists x\,\varphi$	asm, for $\to$I
3	$\quad x_0 \mid \varphi[x_0/x]$	asm, for $\exists$E
4	$\quad\quad \varphi[x_0/x] \to \psi$	$\forall$E 1
5	$\quad\quad \psi$	$\to$E 4, 3
6	$\quad \psi$	$\exists$E 2, 3–5
7	$\exists x\,\varphi \to \psi$	$\to$I 2–6

1	$\exists x\,\varphi \to \psi$	
2	$x_0 \mid \varphi[x_0/x]$	asm, for $\to$I
3	$\quad \exists x\,\varphi$	$\exists$I 2
4	$\quad \psi$	$\to$E 1, 3
5	$\varphi[x_0/x] \to \psi$	$\to$I 2–4
6	$\forall x\,(\varphi \to \psi)$	$\forall$I 2–5

At line 6 on the left, the conclusion ψ does not mention the witness x_0: this is where $\exists$E's side condition meets the condition on ψ.

(c), $\vdash$ and $\dashv$:

1	$\exists x\,(\varphi \wedge \psi)$	
2	$x_0 \mid \varphi[x_0/x] \wedge \psi$	asm, for $\exists$E
3	$\quad \varphi[x_0/x]$	$\wedge$E 2
4	$\quad \exists x\,\varphi$	$\exists$I 3
5	$\quad \psi$	$\wedge$E 2
6	$\quad \exists x\,\varphi \wedge \psi$	$\wedge$I 4, 5
7	$\exists x\,\varphi \wedge \psi$	$\exists$E 1, 2–6

1	$\exists x\,\varphi \wedge \psi$	
2	$\exists x\,\varphi$	$\wedge$E 1
3	ψ	$\wedge$E 1
4	$x_0 \mid \varphi[x_0/x]$	asm, for $\exists$E
5	$\quad \varphi[x_0/x] \wedge \psi$	$\wedge$I 4, 3
6	$\quad \exists x\,(\varphi \wedge \psi)$	$\exists$I 5
7	$\exists x\,(\varphi \wedge \psi)$	$\exists$E 2, 4–6

(d), $\vdash$:

1	$\exists x\,(\varphi \to \psi)$	
2	$\quad \forall x\,\varphi$	asm, for $\to$I
3	$\quad x_0 \mid \varphi[x_0/x] \to \psi$	asm, for $\exists$E
4	$\quad\quad \varphi[x_0/x]$	$\forall$E 2
5	$\quad\quad \psi$	$\to$E 3, 4
6	$\quad \psi$	$\exists$E 1, 3–5
7	$\forall x\,\varphi \to \psi$	$\to$I 2–6

(d), $\dashv$, with LEM on $\forall x\,\varphi$:

1	$\forall x\,\varphi \to \psi$	
2	$\forall x\,\varphi \vee \neg\forall x\,\varphi$	LEM
3	$\quad \forall x\,\varphi$	asm, for $\vee$E
4	$\quad \psi$	$\to$E 1, 3
5	$\quad\quad \varphi$	asm, for $\to$I
6	$\quad\quad \psi$	R 4
7	$\quad \varphi \to \psi$	$\to$I 5–6
8	$\quad \exists x\,(\varphi \to \psi)$	$\exists$I 7
9	$\quad \neg\forall x\,\varphi$	asm, for $\vee$E
10	$\quad \exists x\,\neg\varphi$	QN (a) 9
11	$\quad x_0 \mid \neg\varphi[x_0/x]$	asm, for $\exists$E
12	$\quad\quad \varphi[x_0/x] \to \psi$	VI 11
13	$\quad\quad \exists x\,(\varphi \to \psi)$	$\exists$I 12
14	$\quad \exists x\,(\varphi \to \psi)$	$\exists$E 10, 11–13
15	$\exists x\,(\varphi \to \psi)$	$\vee$E 2, 3–8, 9–14

Line 8 is $\exists$I at the term x itself: $(\varphi \to \psi)[x/x]$ is $\varphi \to \psi$, as in Proposition 5.5.1, and as there the step is sound because every domain is non-empty (Section 7.1). At line 14, x_0 is in neither the conclusion nor line 10, nor in the premise or the open assumption $\neg\forall x\,\varphi$.

(e), $\vdash$ and $\dashv$, the second with LEM on ψ:

1	$\exists x\,(\psi \to \varphi)$	
2	$\quad \psi$	asm, for $\to$I
3	$\quad x_0 \mid \psi \to \varphi[x_0/x]$	asm, for $\exists$E
4	$\quad\quad \varphi[x_0/x]$	$\to$E 3, 2
5	$\quad\quad \exists x\,\varphi$	$\exists$I 4
6	$\quad \exists x\,\varphi$	$\exists$E 1, 3–5
7	$\psi \to \exists x\,\varphi$	$\to$I 2–6

1	$\psi \to \exists x\,\varphi$	
2	$\psi \vee \neg\psi$	LEM
3	$\quad \psi$	asm, for $\vee$E
4	$\quad \exists x\,\varphi$	$\to$E 1, 3
5	$\quad x_0 \mid \varphi[x_0/x]$	asm, for $\exists$E
6	$\quad\quad \psi$	asm, for $\to$I
7	$\quad\quad \varphi[x_0/x]$	R 5
8	$\quad\quad \psi \to \varphi[x_0/x]$	$\to$I 6–7
9	$\quad\quad \exists x\,(\psi \to \varphi)$	$\exists$I 8
10	$\quad \exists x\,(\psi \to \varphi)$	$\exists$E 4, 5–9
11	$\quad \neg\psi$	asm, for $\vee$E
12	$\quad \psi \to \varphi$	VI 11
13	$\quad \exists x\,(\psi \to \varphi)$	$\exists$I 12
14	$\exists x\,(\psi \to \varphi)$	$\vee$E 2, 3–10, 11–13

Line 13 is again $\exists$I at the variable x. At line 10, x_0 is in neither the conclusion $\exists x\,(\psi \to \varphi)$ nor line 4, $\exists x\,\varphi$, nor in the premise or the open assumption ψ of line 3. In each of the eight other derivations, the side conditions are checked as in Proposition 6.2.1, and LEM and RAA do not occur. This proves Proposition 6.2.7.

■

Why the condition on ψ. Drop it, and (b) fails. Take φ and ψ both to be $P(x)$. Then $\forall x\,(P(x) \to P(x))$ is true in every domain, while $\exists x\,P(x) \to P(x)$, over $\mathbb{Z}$ with P read as "is even" and x read as 1, says "if some integer is even, then 1 is even", which is false. The derivation of (b), in the fold above, shows where the condition was used. Without it, line 4 would be $P(x_0) \to P(x_0)$, line 5 would be $P(x_0)$, and the $\exists$E box could not close on a line that mentions its witness.

Second reading · *Passage laws on a part*

A passage law of Proposition 6.2.7 asks that its moving variable y not be free in ψ; since y appears in β, the fresh x_0 is not y. If y is x itself, substitution stops at the quantifier on x, and $\psi[x_0/x]$ is ψ because x is not free in ψ; the pair is then the law's own instance, unchanged. Otherwise substituting x_0 for x passes the quantifier on y unchanged, and $\psi[x_0/x]$ has no free variable that ψ lacks except x_0, and y is not free in it either. So the pair is again an instance of the law, with its condition met.

Quick check. A student writes $\forall x\,(x > 0 \to \exists y\,(y \cdot y = x))$ as $\forall x\,\exists y\,(x > 0 \to y \cdot y = x)$, citing a passage law. Which law, and is the condition met?

Solution. Law (e) of Proposition 6.2.7, applied under $\forall x$ (a law may be used on a part under a quantifier: Proposition 6.2.4, with the fold "Passage laws on a part" above), with ψ the formula $x > 0$, φ the formula $y \cdot y = x$, and y as the moving variable. The condition asks that y not be free in ψ, and y does not occur in $x > 0$; so the step is correct. It is the classical direction of (e), the one derived with LEM.

Second reading · *Twin laws*

Over a domain with two elements a, b, a $\forall$ is an "and" and an $\exists$ is an "or" (Section 5.5 read $\exists$ over a finite domain as a long "or"), and each law of this section becomes a propositional law. Below, φ_a abbreviates $\varphi[a/x]$, the instance at a; a formula ψ in which x is not free is the same at a and at b.

law	over $\{a, b\}$	the propositional twin
QN (a), (b)	$\neg(\varphi_a \wedge \varphi_b) \dashv\vdash \neg\varphi_a \vee \neg\varphi_b$	Proposition 2.5.5 (d), (c)
QN (c)	$\neg(\varphi_a \vee \varphi_b) \dashv\vdash \neg\varphi_a \wedge \neg\varphi_b$	Proposition 2.5.5 (a), (b)
distribution (c)	$(\varphi_a \wedge \varphi_b) \vee (\psi_a \wedge \psi_b) \vdash (\varphi_a \vee \psi_a) \wedge (\varphi_b \vee \psi_b)$	a direct derivation: $\vee$E on the premise, and in each case $\wedge$E twice, $\vee$I twice and $\wedge$I
passage (b)	$(\varphi_a \to \psi) \wedge (\varphi_b \to \psi) \dashv\vdash (\varphi_a \vee \varphi_b) \to \psi$	Example (Cases on a hypothesis are two implications), Example (A disjunctive hypothesis splits into two implications)

In each pair the direction that needs RAA is the same: QN (a) and De Morgan (d). A twin is a reading, not a proof: a domain may be infinite, and the laws were derived for every domain.

6.2.8 Prenex form

A formula is in **prenex form** when all its quantifiers stand at the front, as a **prefix**, before a part with no quantifier, the **matrix**.

Example. Bringing a formula into prenex form

Bring $\exists x\, P(x) \to \exists y\, Q(y)$ into prenex form.

Solution.

$$\begin{aligned} &\exists x\, P(x) \to \exists y\, Q(y) \\ \dashv\vdash\ &\forall x\, (P(x) \to \exists y\, Q(y)) \\ \dashv\vdash\ &\forall x\, \exists y\, (P(x) \to Q(y)). \end{aligned}$$

- The first step is passage law (b), with ψ the formula $\exists y\, Q(y)$, in which x is not free.
- The second is law (e) under $\forall x$, by Proposition 6.2.4, with ψ the formula $P(x)$, in which y is not free. Its hypothesis is checked in the fold "Passage laws on a part".

Going down, the second step uses the classical direction of (e).

Alternation. Count the changes between $\forall$ and $\exists$ along the prefix: the example has one, $\forall x\, \exists y$, and "$a_n \to L$" has two, $\forall \varepsilon\, \exists N\, \forall n$. Each change is one more round of the game of Section 5.1: a challenger move answered by a defender move. A prefix with more alternations takes more rounds to win, and checking it by examples gets harder: no finite table of values settles a $\forall\exists\forall$ sentence over $\mathbb{N}$.

Quick check. How many alternations have the prefixes of "uniformly continuous on A" and "continuous at every point of A" (Section 6.1, Exercise (C1: Where the $\forall x$ stands))?

Solution. Two each: $\forall \varepsilon\, \exists \delta\, \forall x\, \forall y$ and $\forall x\, \forall \varepsilon\, \exists \delta\, \forall y$. The count is the same, and the order differs. Section 6.4 is about that difference.

Remark (Looking ahead). Chapter 8 says what is at stake when "not every" becomes "some not" by RAA, and catalogues the wrong negations of this section among the steps a proof may not take.

6.2.9 Exercises

6.2.9.1 A. Check your understanding

Exercise. A1

These parts are about the derivation of Proposition 6.2.1 (a).

(a) Which lines are justified by RAA, and what does each of the two boxes assume and

conclude?

(b) Why can the derivation not reach $\exists x\, \neg\varphi$ by one $\exists$I directly from the premise?

(c) Line 4 is an $\exists$I at x_0. Why is its conclusion not the conclusion of the derivation, although the two formulas are the same?

(d) Run the derivation with φ the formula $x \cdot x \neq 2$, over $\mathbb{R}$. Write lines 1 and 9 in words. Does any line name the number $\sqrt{2}$? What more would be needed to reach $\exists x\, (x \cdot x = 2)$?

(e) True or false, with a reason: $\exists \varepsilon \leq 0\ \neg\varphi$ is a correct way to write the negation of $\forall \varepsilon > 0\ \varphi$.

Solution. *(a)* Lines 6 and 9. The inner box (lines 3–5) assumes $\neg\varphi[x_0/x]$, reaches $\bot$, and line 6 concludes $\varphi[x_0/x]$. The outer box (lines 2–8) assumes $\neg\exists x\, \neg\varphi$, reaches $\bot$, and line 9 concludes $\exists x\, \neg\varphi$.

(b) $\exists$I needs a line holding an instance $\neg\varphi[t/x]$. The premise $\neg\forall x\, \varphi$ is not such an instance (its main symbol is $\neg$ followed by $\forall$), and no rule takes it apart except $\neg$E, which needs $\forall x\, \varphi$. The premise says a counterexample exists without naming one.

(c) Line 4 lies inside three boxes: the box opened at line 2, and the flagged box of x_0 and the inner box, both opened at line 3. It depends on the assumptions on lines 2 and 3, so it is not available at the end. It is used only to clash with line 2. The conclusion on line 9 is reached by closing the outer box with RAA.

(d) Line 1: "not every real number has a square different from 2". Line 9: "some real number does not have a square different from 2", that is, $\exists x\, \neg(x \cdot x \neq 2)$. No line names $\sqrt{2}$: the only witness on any line is the fresh name x_0, inside boxes that closed, and line 9 comes from RAA, not from $\exists$I. To reach $\exists x\, (x \cdot x = 2)$, open an $\exists$E box on line 9 with the assumption $\neg(x_1 \cdot x_1 \neq 2)$, apply $\neg\neg$E (the formula $x_1 \cdot x_1 \neq 2$ is $\neg(x_1 \cdot x_1 = 2)$) to get $x_1 \cdot x_1 = 2$, and close with $\exists$I.

(e) False. By Proposition 6.2.3 the negation is $\exists \varepsilon > 0\ \neg\varphi$: the bound stays. The wrong form can be true together with the statement it claims to negate. Over $\mathbb{R}$, take φ to be $\varepsilon > 0$: then $\forall \varepsilon > 0\, (\varepsilon > 0)$ is true, and $\exists \varepsilon \leq 0\ \neg(\varepsilon > 0)$ is true as well, with the witness $\varepsilon = 0$. A statement and its negation are never both true.

Exercise. A2: Drill

Each part takes a line or two. In (a)–(h), negate the statement into positive form (in (g), choose the negation); the last four parts use the square, the reading over two elements, prenex form and a law under a quantifier.

(a) $\forall x\, P(x)$.

(b) $\exists x\, (P(x) \wedge Q(x))$.

(c) $\forall x > 0\ (x^2 > 0)$, over $\mathbb{R}$.

(d) "If n is even, then n^2 is even", for one integer n.

(e) $\forall \varepsilon > 0\ \exists \delta > 0\ (\delta < \varepsilon)$.

(f) $\forall x \in A\ (x \leq s)$.

(g) Which is the negation of $\forall x \in A\ \exists y \in B\ (x < y)$: (i) $\exists x \in A\ \forall y \in B\ (x \geq y)$; (ii)

$\exists x \notin A\, \forall y \notin B\, (x \geq y)$; (iii) $\forall x \in A\, \exists y \in B\, (x \geq y)$?

(h) "(a_n) is bounded below": $\exists m\, \forall n \in \mathbb{N}\, (m \leq a_n)$.

(i) Over $\mathbb{N}$, with S "is a multiple of 3" and P "is odd", which corner of the square is the negation of "some multiple of 3 is not odd", and is it true?

(j) Over $\{a, b\}$, write $\exists x\, (P(x) \wedge Q(x))$ and $\exists x\, P(x) \wedge \exists x\, Q(x)$ without quantifiers, and give truth values that make the second true and the first false.

(k) Name the passage law, and give a prenex form of $\exists x\, P(x) \to R$, where x is not free in R.

(l) Is this step correct: from $\forall x\, (P(x) \to Q(x))$ to $\forall x\, (\neg P(x) \to \neg Q(x))$, "contrapositive under $\forall$"?

Solution.

(a) $\exists x\, \neg P(x)$.

(b) $\forall x\, (\neg P(x) \vee \neg Q(x))$, equivalently $\forall x\, (P(x) \to \neg Q(x))$.

(c) $\exists x > 0\, (x^2 \leq 0)$: the bound stays, and the order fact turns $\neg(x^2 > 0)$ into $x^2 \leq 0$.

(d) "n is even and n^2 is not even" (Proposition 6.2.2).

(e) $\exists \varepsilon > 0\, \forall \delta > 0\, (\delta \geq \varepsilon)$, by the order fact for $\neg(\delta < \varepsilon)$.

(f) $\exists x \in A\, (x > s)$.

(g) (i): flip each quantifier, keep each bound.

(h) $\forall m\, \exists n \in \mathbb{N}\, (a_n < m)$, by the order fact for $\neg(m \leq a_n)$.

(i) A, "every multiple of 3 is odd", by the diagonal. It is false: 6 is a multiple of 3 and even.

(j) $(P(a) \wedge Q(a)) \vee (P(b) \wedge Q(b))$ and $(P(a) \vee P(b)) \wedge (Q(a) \vee Q(b))$. Take $P(a)$ and $Q(b)$ true, $P(b)$ and $Q(a)$ false.

(k) Law (b): $\forall x\, (P(x) \to R)$.

(l) No. The contrapositive is $\forall x\, (\neg Q(x) \to \neg P(x))$. Over $\mathbb{Z}$, with P "is a multiple of 4" and Q "is even", the premise is true and the line is false at 2.

6.2.9.2 B. Practice

Exercise. B1: Four negations

Negate each definition into positive form, citing the law used at each step. Work over $\mathbb{R}$, with (a_n) a sequence and f a function.

(a) (a_n) is bounded: $\exists M\, \forall n \in \mathbb{N}\, (|a_n| \leq M)$.

(b) f is increasing on A: $\forall x \in A\, \forall y \in A\, (x < y \to f(x) \leq f(y))$.

(c) $\lim_{x \to a} f(x) = L$: $\forall \varepsilon > 0\, \exists \delta > 0\, \forall x\, ((0 < |x - a| \wedge |x - a| < \delta) \to |f(x) - L| < \varepsilon)$.

(d) (a_n) is a Cauchy sequence: $\forall \varepsilon > 0\, \exists N \in \mathbb{N}\, \forall m \in \mathbb{N}\, \forall n \in \mathbb{N}\, ((m \geq N \wedge n \geq N) \to |a_m - a_n| < \varepsilon)$.

Solution. *(a)* By Proposition 6.2.1 (c) (the $\exists M$ has no bound), then Proposition 6.2.3 (a)

under $\forall M$:

$$\neg\exists M\ \forall n \in \mathbb{N}\ (|a_n| \le M) \dashv\vdash \forall M\ \neg\forall n \in \mathbb{N}\ (|a_n| \le M) \dashv\vdash \forall M\ \exists n \in \mathbb{N}\ \neg(|a_n| \le M),$$

and the order fact turns $\neg(|a_n| \le M)$ into $|a_n| > M$:

$$\forall M\ \exists n \in \mathbb{N}\ (|a_n| > M).$$

Every real number is exceeded by the size of some term.

(b) Two uses of Proposition 6.2.3, then Proposition 6.2.2 and the order fact:

$$\exists x \in A\ \exists y \in A\ (x < y \wedge f(x) > f(y)).$$

Two points of A, in increasing order, whose values are in decreasing order.

(c) Proposition 6.2.3 twice, Proposition 6.2.1 (a) and (b) for the unbounded $\forall x$, Proposition 6.2.2, and the order fact:

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x\ ((0 < |x - a| \wedge |x - a| < \delta) \wedge |f(x) - L| \ge \varepsilon).$$

The hypothesis $0 < |x - a| < \delta$ is not negated: it moves, unchanged, into a conjunction.

(d) Proposition 6.2.3 four times, Proposition 6.2.2, the order fact:

$$\exists \varepsilon > 0\ \forall N \in \mathbb{N}\ \exists m \in \mathbb{N}\ \exists n \in \mathbb{N}\ ((m \ge N \wedge n \ge N) \wedge |a_m - a_n| \ge \varepsilon).$$

For some $\varepsilon > 0$, however far out we start, two later terms are at least ε apart.

Exercise. B2: Two passage laws in one derivation

Derive $\exists x\, P(x) \to \forall y\, Q(y) \vdash \forall x\, \forall y\, (P(x) \to Q(y))$ without RAA. Then name the two laws of Proposition 6.2.7 that the sequent combines, and say how.

Solution. The goal is $\forall x$, then $\forall y$, then $\to$: open two flagged boxes and a $\to$I box, in the goal's order. Inside, the premise is an implication whose hypothesis $\exists x\, P(x)$ comes from the assumption $P(x_0)$ by $\exists$I.

1	$\exists x\, P(x) \to \forall y\, Q(y)$				
2	x_0	y_0		$P(x_0)$	asm, for $\to$I
3				$\exists x\, P(x)$	$\exists$I 2
4				$\forall y\, Q(y)$	$\to$E 1, 3
5				$Q(y_0)$	$\forall$E 4
6			$P(x_0) \to Q(y_0)$		$\to$I 2–5
7		$\forall y\, (P(x_0) \to Q(y))$			$\forall$I 2–6
8	$\forall x\, \forall y\, (P(x) \to Q(y))$				$\forall$I 2–7

At line 7, y_0 is not in the premise or in line 7, and no assumption is open (the box of line 2

has closed). At line 8 the same holds for x_0. No RAA is used.

The sequent combines (b) and (a) of Proposition 6.2.7. Law (b), with φ the formula $P(x)$ and ψ the formula $\forall y\, Q(y)$ (in which x is not free), turns the premise into $\forall x\,(P(x) \to \forall y\, Q(y))$. Then law (a), applied under $\forall x$ (a provable equivalence may be used under a quantifier, by Proposition 6.2.4) with ψ the formula $P(x)$, φ the formula $Q(y)$ and y as the moving variable (y is not free in $P(x)$), turns $P(x) \to \forall y\, Q(y)$ into $\forall y\,(P(x) \to Q(y))$.

Exercise. B3: Choosing ε from the data

Let a and b be real numbers with $a < b + \varepsilon^2$ for every $\varepsilon > 0$. Prove that $a \le b$, by the method of Proposition 6.2.5: suppose not, and choose an ε built from the data that makes the numbers collide. Then explain why the choice $\varepsilon = a - b$ does not work.

Solution. Suppose $a < b + \varepsilon^2$ for every $\varepsilon > 0$, and suppose, for a contradiction, that $a > b$ (the negation of $a \le b$). Let $\varepsilon := \min(1, a - b)$. Then $\varepsilon > 0$, since $1 > 0$ and $a - b > 0$. Since $0 < \varepsilon \le 1$, $\varepsilon^2 = \varepsilon \cdot \varepsilon \le \varepsilon \cdot 1 = \varepsilon$, and $\varepsilon \le a - b$. The hypothesis applies to this ε and gives

$$a < b + \varepsilon^2 \le b + \varepsilon \le b + (a - b) = a,$$

so $a < a$, a contradiction. Hence $a \le b$.

The choice $\varepsilon = a - b$, copied from the text, gives $a < b + (a - b)^2$, that is, $a - b < (a - b)^2$. When $a - b > 1$ this is true (for $a - b = 2$: $2 < 4$), so there is no contradiction. The square shrinks a number only when it is at most 1, and the min with 1 is what puts ε there.

Exercise. B4: Derive or break

Derive each claim with the rules and laws so far, or break it by a reading over $\mathbb{Z}$ that makes every member of its counterexample set true.

(a) $\forall x\,(P(x) \leftrightarrow Q(x)) \vdash \forall x\, P(x) \leftrightarrow \forall x\, Q(x)$.
(b) $\forall x\, P(x) \leftrightarrow \forall x\, Q(x) \vdash \forall x\,(P(x) \leftrightarrow Q(x))$.
(c) $\neg\exists x\,(P(x) \wedge \neg Q(x)) \vdash \forall x\,(P(x) \to Q(x))$.
(d) $\exists x\, P(x) \to \exists x\, Q(x) \vdash \forall x\,(P(x) \to Q(x))$.
(e) $\forall x\,(P(x) \vee Q(x)) \vdash \forall x\, P(x) \vee \exists x\, Q(x)$.
(f) $\forall x\, P(x) \vee \exists x\, Q(x) \vdash \forall x\,(P(x) \vee Q(x))$.

Solution.

(a) Derive. The goal is a biconditional, so $\leftrightarrow$ I closes two boxes (Definition 2.4.1); in each, the goal is a $\forall$, so a flagged box.

1		$\forall x\,(P(x) \leftrightarrow Q(x))$	
2		$\forall x\,P(x)$	asm, for $\leftrightarrow$I
3	x_0	$P(x_0) \leftrightarrow Q(x_0)$	$\forall$E 1
4		$P(x_0)$	$\forall$E 2
5		$Q(x_0)$	$\leftrightarrow$E 3, 4
6		$\forall x\,Q(x)$	$\forall$I 3–5
7		$\forall x\,Q(x)$	asm, for $\leftrightarrow$I
8	x_1	$P(x_1) \leftrightarrow Q(x_1)$	$\forall$E 1
9		$Q(x_1)$	$\forall$E 7
10		$P(x_1)$	$\leftrightarrow$E 8, 9
11		$\forall x\,P(x)$	$\forall$I 8–10
12		$\forall x\,P(x) \leftrightarrow \forall x\,Q(x)$	$\leftrightarrow$I 2–6, 7–11

At line 6, x_0 is in neither the premise, the open assumption on line 2, nor $\forall x\,Q(x)$; line 11 is checked the same way for x_1.

(b) Breaks. Under the reading with P "is even" and Q "is odd", every member of the counterexample set is true. Both $\forall x\,P(x)$ and $\forall x\,Q(x)$ are false, so the premise is true; at 0, "0 is even if and only if 0 is odd" is false, so the negated conclusion is true.

(c) Derive. Turn the premise into a $\forall$ by QN (c), then open a flagged box for the goal.

1		$\neg\exists x\,(P(x) \wedge \neg Q(x))$	
2		$\forall x\,\neg(P(x) \wedge \neg Q(x))$	QN (c) 1
3	x_0	$\neg(P(x_0) \wedge \neg Q(x_0))$	$\forall$E 2
4		$P(x_0)$	asm, for $\rightarrow$I
5		$\neg Q(x_0)$	asm, for RAA
6		$P(x_0) \wedge \neg Q(x_0)$	$\wedge$I 4, 5
7		$\bot$	$\neg$E 6, 3
8		$Q(x_0)$	RAA 5–7
9		$P(x_0) \rightarrow Q(x_0)$	$\rightarrow$I 4–8
10		$\forall x\,(P(x) \rightarrow Q(x))$	$\forall$I 3–9

At line 10, x_0 is in neither the premise nor the conclusion, and no assumption is open.

(d) Breaks. Under the reading with P "is even" and Q "is positive", every member of the counterexample set is true. Some integer is even (2) and some integer is positive (1), so the premise is true; "every even integer is positive" fails at 0, so the negated conclusion is true.

(e) Derive. Decide $\forall x\,P(x)$ by LEM; in the second case QN (a) supplies a witness that fails P, and the premise gives it Q.

1	$\forall x\,(P(x) \lor Q(x))$	
2	$\forall x\,P(x) \lor \neg\forall x\,P(x)$	LEM
3	$\quad \forall x\,P(x)$	asm, for $\lor$E
4	$\quad \forall x\,P(x) \lor \exists x\,Q(x)$	$\lor$I 3
5	$\quad \neg\forall x\,P(x)$	asm, for $\lor$E
6	$\quad \exists x\,\neg P(x)$	QN (a) 5
7	$\quad x_0 \mid \neg P(x_0)$	asm, for $\exists$E
8	$\quad\quad P(x_0) \lor Q(x_0)$	$\forall$E 1
9	$\quad\quad Q(x_0)$	DS 8, 7
10	$\quad\quad \exists x\,Q(x)$	$\exists$I 9
11	$\quad \exists x\,Q(x)$	$\exists$E 6, 7–10
12	$\quad \forall x\,P(x) \lor \exists x\,Q(x)$	$\lor$I 11
13	$\forall x\,P(x) \lor \exists x\,Q(x)$	$\lor$E 2, 3–4, 5–12

At line 11, x_0 is not in line 11, in line 6, in the premise or in the open assumption on line 5.

(f) Breaks. Under the reading with P "is even" and $Q(x)$ "$x = 0$", every member of the counterexample set is true. Some integer is 0, so $\exists x\,Q(x)$ and the premise are true; "every integer is even or equal to 0" fails at 1, so the negated conclusion is true.

(g) and (b), and (e) and (f), are converse pairs; (d) asks for more than its premise gives. The breaks are counter-readings (Section 7.2).

Marking scheme. (10 points)

- 6: one for each correct verdict.
- 2: the derivations of (a), (c) and (e), with the side condition checked at every closing line of a flagged or $\exists$E box.
- 2: the three readings, each with the truth values of the premise and of the conclusion justified.

Fatal errors (0 for the part): a "break" whose reading leaves the premise false; a derivation that closes an $\exists$E box on a line that mentions its witness.

Exercise. B5: Find the error

Here is a "derivation" of $\exists x\,(P(x) \land Q(x)),\ \forall x\,(Q(x) \to R(x)) \vdash \exists x\,R(x)$.

1	$\exists x\,(P(x) \wedge Q(x))$	
2	$\forall x\,(Q(x) \to R(x))$	
3	$P(c) \wedge Q(c)$	$\exists$E 1
4	$Q(c)$	$\wedge$E 3
5	$Q(c) \to R(c)$	$\forall$E 2
6	$R(c)$	$\to$E 5, 4
7	$\exists x\,R(x)$	$\exists$I 6

Is the conclusion derivable? Does this derivation prove it? Find the line that fails, break the step that line takes, and repair the derivation.

Solution. The conclusion is derivable, and this derivation does not prove it. Line 3 fails: $\exists$E is not a one-line step. It opens a box whose witness is a fresh variable, never a constant (a constant has a fixed meaning), and the box must close on a line that does not mention the witness; here no box is opened. Nothing licenses passing from "something has P and Q" to "c has them".

The step from $\exists x\,(P(x) \wedge Q(x))$ to $P(c) \wedge Q(c)$ breaks over $\mathbb{Z}$, with P "is even", Q "is positive" and c read as 0: the premise is true (2 is both), and the line is false, since 0 is not positive.

The repair opens an $\exists$E box:

1		$\exists x\,(P(x) \wedge Q(x))$	
2		$\forall x\,(Q(x) \to R(x))$	
3	x_0	$P(x_0) \wedge Q(x_0)$	asm, for $\exists$E
4	x_0	$Q(x_0)$	$\wedge$E 3
5	x_0	$Q(x_0) \to R(x_0)$	$\forall$E 2
6	x_0	$R(x_0)$	$\to$E 5, 4
7	x_0	$\exists x\,R(x)$	$\exists$I 6
8		$\exists x\,R(x)$	$\exists$E 1, 3–7

At line 8, x_0 is not in line 8, in line 1 or in a premise, and no other assumption is open.

Marking scheme. (10 points)

- 2: "derivable, but not proved by this derivation".
- 3: line 3, and the rule it breaks: $\exists$E needs a box with a fresh witness.
- 2: a reading that breaks the step, with both truth values.
- 3: a repaired derivation, with the side condition of $\exists$E checked.

Fatal error (0 for the step): naming line 7 ($\exists$I is legal there).

6.2.9.3 C. Going deeper

Exercise. C1: Two prenex forms

A formula is in **prenex form** when all its quantifiers stand at the front, before a part with no quantifier. Using the laws of Proposition 6.2.7, bring $\forall x\, P(x) \to \exists y\, Q(y)$ into prenex form in two ways, ending with $\exists x\, \exists y\, (P(x) \to Q(y))$ and with $\exists y\, \exists x\, (P(x) \to Q(y))$. Justify each step, and say which directions of the laws are classical.

Solution (Hint 1). The formula is an implication with a quantifier on each side. Law (d) moves the quantifier out of the hypothesis, and law (e) moves it out of the conclusion.

Solution (Hint 2). For the first route, start with law (d), taking ψ to be the whole conclusion $\exists y\, Q(y)$.

Solution. First route: x out first. Law (d) of Proposition 6.2.7, with φ the formula $P(x)$ and ψ the formula $\exists y\, Q(y)$, in which x is not free:

$$\forall x\, P(x) \to \exists y\, Q(y) \;\dashv\vdash\; \exists x\, (P(x) \to \exists y\, Q(y)).$$

Then law (e) under $\exists x$, with ψ the formula $P(x)$, φ the formula $Q(y)$, and y as the moving variable (y is not free in $P(x)$): $P(x) \to \exists y\, Q(y) \;\dashv\vdash\; \exists y\, (P(x) \to Q(y))$. A provable equivalence may be used under a quantifier (Proposition 6.2.4), so

$$\exists x\, (P(x) \to \exists y\, Q(y)) \;\dashv\vdash\; \exists x\, \exists y\, (P(x) \to Q(y)).$$

Second route: y out first. Law (e), with ψ the formula $\forall x\, P(x)$, φ the formula $Q(y)$, and y as the moving variable (y is not free in $\forall x\, P(x)$):

$$\forall x\, P(x) \to \exists y\, Q(y) \;\dashv\vdash\; \exists y\, (\forall x\, P(x) \to Q(y)).$$

Then law (d) under $\exists y$, with φ the formula $P(x)$ and ψ the formula $Q(y)$ (x is not free in $Q(y)$):

$$\exists y\, (\forall x\, P(x) \to Q(y)) \;\dashv\vdash\; \exists y\, \exists x\, (P(x) \to Q(y)).$$

Classical steps. Going from the given formula towards the prenex form, every step of both routes uses the direction $\dashv$ of (d) or of (e): the direction from the implication to the $\exists$. These are exactly the two directions that Proposition 6.2.7 derives with LEM. The way back uses only the intuitionistic directions.

Marking scheme. (10 points) - 3: first route, with law (d) then law (e), each with φ, ψ and the moving variable named, and the "not free" condition checked. - 3: second route, likewise with (e) then (d). - 2: says why a law may be applied under $\exists x$ or $\exists y$ (an equivalence used under a quantifier, by an $\exists$E box closed with $\exists$I). - 2: identifies the $\dashv$ directions of (d) and (e) as the classical ones, used in every step towards prenex form. Fatal errors (0 for the step): moving $\forall x$ out of the hypothesis as $\forall x$ (it becomes $\exists x$); applying a passage law when the moving variable is free in ψ.

Exercise. C2: Exit check: negations

(a) A sequence of functions f_n converges to f **uniformly** on A when

$$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ \forall x \in A\ (n \geq N \to |f_n(x) - f(x)| < \varepsilon),$$

where $f_n(x)$ is a term in the two variables n and x. Negate this into positive form, citing the laws.

(b) Using the negation of uniform continuity in Example (Three negations from analysis) (c), prove that $f(x) = 1/x$ is not uniformly continuous on $(0, 1)$. Give the witnesses that the negation asks for, each as a function of what is chosen before it, and check every condition.

Solution (Hint 1). For (b), the negation asks for one ε first. Near 0, the graph of $1/x$ is steep: two points close together have values far apart.

Solution (Hint 2). Take $\varepsilon = 1$ and let $\delta > 0$ be arbitrary. You must now find $x, y \in (0, 1)$, depending on δ, with $|x - y| < \delta$ and $|1/x - 1/y| \geq 1$.

Solution. *(a)* Proposition 6.2.3 four times, then Proposition 6.2.2 and the order fact:

$$\exists \varepsilon > 0\ \forall N \in \mathbb{N}\ \exists n \in \mathbb{N}\ \exists x \in A\ (n \geq N \wedge |f_n(x) - f(x)| \geq \varepsilon).$$

In words: for some $\varepsilon > 0$, however far out we start, some later f_n is at least ε away from f at some point of A. The point may change with n.

(b) By Example (Three negations from analysis) (c), we must prove

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x \in (0,1)\ \exists y \in (0,1)\ \left(|x - y| < \delta \wedge \left|\frac{1}{x} - \frac{1}{y}\right| \geq \varepsilon\right).$$

Take $\varepsilon = 1$, which is positive. Let $\delta > 0$ be arbitrary, and let $t := \min(\delta, \frac{1}{2})$, so $0 < t \leq \frac{1}{2}$ and $t \leq \delta$; then take $x := t$ and $y := t/2$, which depend on δ. Check:

- $x, y \in (0, 1)$: $0 < t/2 < t \leq \frac{1}{2} < 1$.
- $|x - y| = t/2 < t \leq \delta$, since $t > 0$.
- $|1/x - 1/y| = |1/t - 2/t| = 1/t \geq 2 \geq 1 = \varepsilon$, since $t \leq \frac{1}{2}$.

Since $\delta > 0$ was arbitrary, the negation of uniform continuity holds, so $1/x$ is not uniformly continuous on $(0, 1)$.

Marking scheme. (10 points) - 3: (a) the negation, with every bound kept, $\wedge$ in place of $\to$, and the quantifiers in the order $\exists \varepsilon\ \forall N\ \exists n\ \exists x \in A$. - 2: (b) the witness ε, stated before δ is given. - 2: for an arbitrary $\delta > 0$, points x and y given as functions of δ. - 3: the three checks: $x, y \in (0, 1)$; $|x - y| < \delta$; $|f(x) - f(y)| \geq \varepsilon$. A different correct choice earns full marks (for instance $x = t$, $y = t/3$); the scheme rewards the order of the choices and the checks. Fatal errors (0 for the step): ε chosen depending on δ; points chosen before δ is given, or not depending on it where they must; a bound negated instead of the quantifier ($\exists \varepsilon \leq 0$, $\exists n < N$).

Exercise. C3: "There exists" from "for every"

Suppose $\exists$ were not a symbol of the language, and $\exists x\,\varphi$ abbreviated $\neg\forall x\,\neg\varphi$.

(a) Show that $\exists$I is then a derived step: from $\varphi[t/x]$, with t free for x in φ, derive $\neg\forall x\,\neg\varphi$.

(b) Show that $\exists$E is one too: if $\Gamma, \varphi[x_0/x] \vdash \chi$, where x_0 is a variable other than x that appears neither in Γ, nor in χ, nor in φ, then $\Gamma, \neg\forall x\,\neg\varphi \vdash \chi$.

(c) Which of (a) and (b) uses a classical rule (RAA or LEM)?

Solution (Hint 1). In (a) the goal is a negation: open a $\neg$I box.

Solution (Hint 2). In (b) nothing in the premises can be taken apart; aim the clash at the premise $\neg\forall x\,\neg\varphi$.

Solution. (a) Five lines:

1	$\varphi[t/x]$	
2	$\quad\forall x\,\neg\varphi$	asm, for $\neg$I
3	$\quad\neg\varphi[t/x]$	$\forall$E 2
4	$\quad\bot$	$\neg$E 1, 3
5	$\neg\forall x\,\neg\varphi$	$\neg$I 2–4

Line 3 is $\forall$E at t, which is free for x in $\neg\varphi$ because it is free for x in φ.

(b) From the premises Γ and $\neg\forall x\,\neg\varphi$:

- assume $\neg\chi$, for RAA;
- open a flagged box x_0, and in it assume $\varphi[x_0/x]$, for $\neg$I;
- paste the given derivation of χ, whose premises Γ are available. Its own flags are first renamed to names that appear nowhere else, as "The laws as one line" requires, so that the open assumption $\neg\chi$ cannot spoil its side conditions;
- write $\bot$ by $\neg$E with $\neg\chi$, and $\neg\varphi[x_0/x]$ by $\neg$I; close the flagged box with $\forall x\,\neg\varphi$ by $\forall$I. Here x_0 is in no premise, not in the open assumption $\neg\chi$ and not in $\forall x\,\neg\varphi$;
- write $\bot$ by $\neg$E with the premise $\neg\forall x\,\neg\varphi$, and χ by RAA.

(c) (b) uses one: RAA at its last line (or LEM, in a proof by cases); (a) uses neither. (Whether (b) could avoid a classical rule altogether is not asked.)

Marking scheme. (10 points)

- 3: (a).
- 2: (b)'s classical step (RAA, or LEM with a case split) and flagged box, in the right nesting.
- 3: the side condition of the $\forall$I line, including the renaming of the pasted flags.
- 2: (c).

Fatal errors (0 for the step): $\forall$I on x_0 while $\varphi[x_0/x]$ is still open; instantiating $\neg\forall x\,\neg\varphi$ by

∀E.

6.3 Finding First-Order Derivations

Chapter 5 added one step to the search method in each rule section, and Section 6.2 added laws. This section puts the steps into one method, puts in order the moves that do most of the work with quantifiers and names a third, admits the propositional step TAUT into derivations with quantifiers, and runs the four kinds of task of Section 3.2 again. The new difficulty is names: which term to instantiate at, when to open a box for a witness, and which names a line may mention when a box closes.

Remark (Looking back). Section 3.1 assembled the search method for the connectives, and Section 3.2 sorted the tasks it serves. Here both return with flags and witnesses. Backward from the goal and forward from the premises is unchanged; what is new is that a box may carry a name, and a closing line must not mention it.

6.3.1 The search method, extended

Section 3.1 put the moves for the connectives in order. Each of Sections 5.3–5.7 added a step. Here is the method with all of them.

Algorithm. **Searching for a derivation of $\Gamma \vdash \varphi$ in predicate logic.**

1. **Set up.** Write the premises Γ at the top and φ at the bottom, with one gap between. Translate any definition in words into a formula first, as in Example (Five definitions, translated).
2. **Goal-directed step.** Take the current goal's main symbol and set up its introduction rule:
   - $\wedge$, $\to$, $\neg$, $\leftrightarrow$: as in Section 3.1;
   - $\forall x\,\psi$: open a **flagged box** with a fresh name x_0; the new goal, the box's last line, is $\psi[x_0/x]$ (Section 5.4). For $\forall x \in A\ \psi$, the box opens with the assumption $x_0 \in A$;
   - $\vee$: **postpone**, as in Section 3.1;
   - $\exists x\,\psi$: **postpone**; work on the premises until an available line shows a term t with $\psi[t/x]$, then use $\exists$I (Section 5.5). For $\exists x \in A\ \psi$, the line $t \in A$ is needed too.

   Repeat while the goal is $\wedge$, $\to$, $\neg$, $\leftrightarrow$, or $\forall$.
3. **Premise-directed step.** Take the main symbol of a resource and apply its elimination rule:
   - $\wedge$, $\to$, $\vee$, $\leftrightarrow$, and the derived rules: as in Section 3.1;
   - $\exists x\,\psi$: open the $\exists$E box **early**, as soon as the line is available, with a fresh witness x_0, and make its last line the current goal, which must not mention x_0 (Section 5.6). An early box lets the witness serve the goals that come later;
   - $\forall x\,\psi$: do not instantiate blindly; instantiate at the term that the goal, or another resource, mentions (Section 5.3). For $\forall x \in A\ \psi$, the line $t \in A$ is needed too;
   - an equation $t_1 = t_2$: use =E with a φ chosen so that one substitution turns an available line into the goal, turning the equation round first if needed (Section 5.7).
4. **Clash.** As in Section 3.1: for the goal $\bot$, find or manufacture a pair χ, $\neg\chi$; an

impossible case closes itself by ⊥E.

5. **Stuck.** As in Section 3.1: RAA on the current goal, or LEM on a well-chosen statement.

When no gap is left, number the lines and check the derivation as Definition 1.4.2, extended in Section 5.4, says: every citation is available, and every side condition holds (t free for x at ∀E and ∃I, bounded or not; t_1 and t_2 free for x at =E; a fresh name at each line that closes a flagged box).

Two of the new steps deserve a word on their place in the order. The goal step for ∀ is safe in the same way as the goal steps for → and ¬ (Section 3.1): if $\forall x\,\psi$ is derivable at all, then so is $\psi[x_0/x]$ for a fresh x_0, by one more ∀E; so opening the box loses nothing. The goal step for ∃ is unsafe in the same way as the one for ∨: a witness chosen too early may be the wrong one, as in Example (Derivations with existential introduction), where "take $y = c$" stalled. That is why ∃ is postponed and the ∃E boxes of the premises are opened early: the witnesses they supply are often the terms the postponed goal is waiting for, as N_0 was in the limit derivation of Section 6.1.

6.3.2 Three moves for quantifiers

Most first-order searches turn on three moves. The first two carry the names Chapter 5 gave them; what is new here is their order, which the method above already follows.

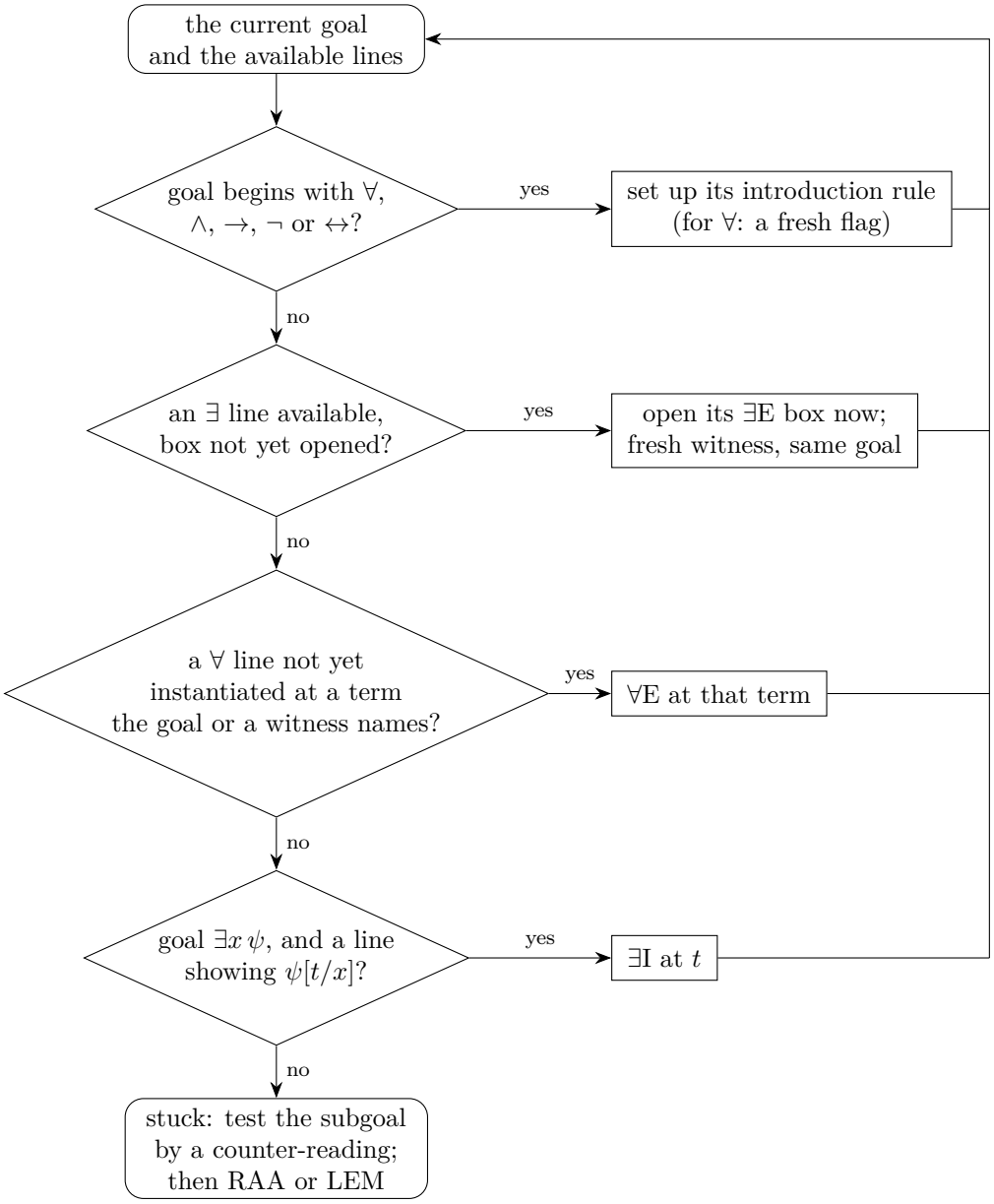
Name the witness early (Section 5.6). When an ∃ line is among the resources, open its box at once, with a fresh witness, and make the box's last line the current goal. The witness costs nothing to name, and it is often the very term that a ∀ line or the goal is waiting for.

Instantiate at the term you need (Section 5.3), and so not before that term is known. A ∀ line can be instantiated at any term at all, which is exactly why instantiating it before the goal or a witness names the term is a blind choice. Wait for the witnesses, then instantiate at them.

The propositional core. Once the boxes are open and the ∀ lines are instantiated, what is left is often a step that holds for propositional reasons alone. The subsection "The propositional core" below names this move and gives it a one-line step.

Two pairs of steps recur: ∀E followed by →E, which applies a general law at one point, and ∃E followed, inside its box, by ∃I with a term built from the names in hand.

The next move, read off the goal and the lines in hand:



The last edge reads "no", not "impossible": the chart is a guide to the next move, and it makes no claim that every derivable sequent is found by following it. It leaves out the premise-directed connective steps, the clash and the propositional core (TAUT, below); for those, follow the method above.

Example. A search that waits for its witnesses

Show that $\forall x\,(P(x) \to \exists y\,R(x,y)),\ \forall x\,\forall y\,(R(x,y) \to Q(y)),\ \exists x\,P(x) \vdash \exists y\,Q(y)$.

Idea. The goal is an ∃: postpone it. Premise 3 is an ∃: open its box now, with witness x_0 (left). Premise 1 is a ∀, and x_0 is now the term it needs: ∀E then →E gives an $\exists y$ line. Open that box at once, with witness y_0 (right). Premise 2 needs two terms, and both are in hand: x_0, then y_0. The postponed goal gets its witness, y_0.

In the left snapshot the gap's goal is still $\exists y\,Q(y)$: an ∃E box keeps the goal it was opened for, and its assumption, line 4, says "$P(x_0)$" and nothing more. In the right snapshot the

same goal has moved one box further in, and line 7 is now the only line that asserts R of particular terms.

1	$\forall x\,(P(x) \to \exists y\,R(x,y))$	
2	$\forall x\,\forall y\,(R(x,y) \to Q(y))$	
3	$\exists x\,P(x)$	
4	x_0 | $P(x_0)$	asm, for $\exists$E
	⋮	
12	$\exists y\,Q(y)$	?
13	$\exists y\,Q(y)$	?

1	$\forall x\,(P(x) \to \exists y\,R(x,y))$	
2	$\forall x\,\forall y\,(R(x,y) \to Q(y))$	
3	$\exists x\,P(x)$	
4	x_0 | $P(x_0)$	asm, for $\exists$E
5	$P(x_0) \to \exists y\,R(x_0,y)$	$\forall$E 1
6	$\exists y\,R(x_0,y)$	$\to$E 5, 4
7	y_0 | $R(x_0,y_0)$	asm, for $\exists$E
	⋮	
11	$\exists y\,Q(y)$	?
12	$\exists y\,Q(y)$	?
13	$\exists y\,Q(y)$	?

Solution.

1		$\forall x\,(P(x) \to \exists y\,R(x,y))$	
2		$\forall x\,\forall y\,(R(x,y) \to Q(y))$	
3		$\exists x\,P(x)$	
4	x_0	$P(x_0)$	asm, for $\exists$E
5		$P(x_0) \to \exists y\,R(x_0,y)$	$\forall$E 1
6		$\exists y\,R(x_0,y)$	$\to$E 5, 4
7		y_0 | $R(x_0,y_0)$	asm, for $\exists$E
8		$\forall y\,(R(x_0,y) \to Q(y))$	$\forall$E 2
9		$R(x_0,y_0) \to Q(y_0)$	$\forall$E 8
10		$Q(y_0)$	$\to$E 9, 7
11		$\exists y\,Q(y)$	$\exists$I 10
12		$\exists y\,Q(y)$	$\exists$E 6, 7–11
13		$\exists y\,Q(y)$	$\exists$E 3, 4–12

At line 12, y_0 is not in line 12, in line 6, in the premises, or in the open assumption on line 4. At line 13, x_0 is not in line 13, in line 3, or in the premises, and no assumption is open.

Read over $\mathbb{Z}$, with $P(x)$ as "$x > 0$", $R(x,y)$ as "$y = 2x + 1$" and $Q(y)$ as "y is odd": every positive integer has a $2x+1$; every number of the form $2x+1$ is odd; some integer is positive; so some integer is odd.

Compare a search that starts with premise 1, before any witness is named. It must instantiate at some term, a variable or a constant that nothing mentions, say c; the line $P(c) \to \exists y\,R(c,y)$ then waits for $P(c)$, which no premise gives. The box of premise 3 supplies the term that makes premise 1 usable.

Quick check. Why is line 8 an instance at x_0, and not at some other term?

Solution. Line 7, the only line that asserts R of particular terms, mentions x_0 and y_0. Premise 2 is useful only at the terms that line names. This is the move "instantiate at the term you need": the terms were known only once the witnesses were named.

6.3.3 Steps that are right for propositional reasons

Inside the boxes of a first-order derivation, many steps are propositional: $\wedge$E, $\to$E, a contrapositive, a case split. Section 4.8 let a propositional derivation take such steps in one line, as TAUT. With quantifiers the step must not look inside a quantifier, and the next definition says exactly what it may look at.

Definition 6.3.1. Propositional Skeleton

Call a first-order formula **prime** if it is atomic and not $\bot$, or begins with a quantifier, bounded or not. Assign to each prime formula a propositional atom, the **same** atom to two prime formulas exactly when one can be turned into the other by finitely many renamings of bound variables (Section 5.2), in either direction; this is an equivalence relation, so the assignment is well defined. The **propositional skeleton** $\mathrm{sk}(\varphi)$ of a formula φ is defined by recursion: $\mathrm{sk}(\bot) = \bot$; $\mathrm{sk}(\pi)$ is the atom of π for a prime formula π; $\mathrm{sk}(\neg\varphi) = \neg\,\mathrm{sk}(\varphi)$; and $\mathrm{sk}(\varphi \circ \psi) = \mathrm{sk}(\varphi) \circ \mathrm{sk}(\psi)$ for each of $\wedge$, $\vee$, $\to$, $\leftrightarrow$. The formulas of one list take their skeletons from one assignment.

The skeleton keeps the connectives outside every quantifier and hides everything else; this is not the skeleton of a derivation that Section 3.1 reads off the goal, and wherever the two could be confused, the book says "propositional skeleton" for this one. For example,

$$\mathrm{sk}(\forall x\, P(x) \to (\exists y\, R(y, c) \vee \neg\forall z\, P(z))) = p \to (q \vee \neg p),$$

since $\forall x\, P(x)$ and $\forall z\, P(z)$ differ only by renaming and so share the atom p.

Why only renaming? The formulas $\forall x\,(P(x) \wedge Q(x))$ and $\forall x\, P(x) \wedge \forall x\, Q(x)$ are provably equivalent (Proposition 6.2.6 (a)), yet the first is prime and the second is a conjunction of two prime formulas. If provable equivalence were allowed to merge atoms, the skeleton would depend on what has been derived, and a TAUT line could hide a quantifier step inside a table. Renaming is different: whether two formulas differ only by renaming bound variables can be seen by comparing them symbol by symbol, so anyone can compute the skeleton without a derivation. And $\bot$ is kept as itself, not given an atom, so that $\neg\bot$ is a tautology and a cited line $\bot$ entails every skeleton, as $\bot$E says it should; an atom would have neither property.

Proposition 6.3.1. Tautological Steps on the Skeleton

Let Γ be a list of first-order formulas and ψ a formula, with skeletons taken from one assignment. If the skeletons of Γ semantically entail $\mathrm{sk}(\psi)$ (Definition 4.5.3), then $\Gamma \vdash \psi$. If no two different prime formulas occurring outside all quantifiers in Γ and ψ share an atom, the derivation uses only the rules of Chapters 1 and 2.

In words: a step that is right for propositional reasons alone is derivable, with the quantified parts carried along untouched.

The proof takes the propositional derivation that completeness provides and puts the prime formulas back in place of their atoms; it is in the fold below.

↻ **Second reading** · *The proof: completeness on the skeleton, then substitution*

Proof. Step 1: no two different prime formulas outside all quantifiers share an atom. By Corollary 4.8.5 (completeness, Section 4.8) there is a derivation D of $\mathrm{sk}(\Gamma) \vdash \mathrm{sk}(\psi)$ with the rules of Chapters 1 and 2; its lines are propositional formulas. Replace, in every line of D, each atom of $\mathrm{sk}(\Gamma)$, $\mathrm{sk}(\psi)$ by the prime formula of Γ, ψ to which it was assigned, and any other atom by $\bot$. This is the substitution of formulas for atoms that Definition 2.3.1 makes for letters, and that Lemma 4.5.3 makes for meaning. Every rule of Chapters 1 and 2 is a schema in formulas, and the replacement commutes with each connective and leaves

$\bot$ alone. So each line keeps its rule and its citations, and each box keeps its extent and its assumption; this holds for RAA, LEM and $\bot$E as for the others. D has no flagged box, so no side condition about names arises. Each premise $\mathrm{sk}(\gamma)$ becomes γ, and the last line becomes ψ. So the result is a first-order derivation of $\Gamma \vdash \psi$ (Definition 1.4.2, as Section 5.4 extended it), and it uses only the rules of Chapters 1 and 2. This is the second sentence of the statement.

Step 2: the general case. Choose one prime formula for each atom, and let γ^* be γ with every prime formula occurring outside all quantifiers (the occurrences the skeleton sees) replaced by the chosen one of its atom. Prime formulas with one atom differ only by renaming. A renaming is a provable equivalence whose closure is derivable (the renaming sentence after Proposition 6.2.4), so Proposition 6.2.4, applied one occurrence and one renaming step at a time (through a letter new to the formula where needed), with the steps chained by Lemma 2.3.3 (c), gives $\gamma \dashv\vdash \gamma^*$ for each formula γ of the list, and $\psi \dashv\vdash \psi^*$. The formulas Γ^* and ψ^* have the same skeletons as Γ and ψ, and in them no two different prime formulas outside all quantifiers share an atom, so Step 1 gives $\Gamma^* \vdash \psi^*$. Now Lemma 2.3.3 (Section 2.3) chains $\Gamma \vdash \gamma^*$ for each γ, then $\Gamma^* \vdash \psi^*$, then $\psi^* \vdash \psi$. Each pasted block's flags are first renamed to names new to the whole derivation, as "The laws as one line" (Section 6.2) argues. Then no flag of a pasted block occurs in a premise or an open assumption of the derivation it is pasted into, so every $\forall$I and $\exists$E side condition of the block still holds, and Lemma 2.3.3 (c) holds for these first-order derivations by the same pasting. So $\Gamma \vdash \psi$.

■

TAUT on the skeleton. From now on, a line of a derivation may be justified TAUT, citing available lines, when their skeletons semantically entail its skeleton (Definition 4.5.3), all from one assignment; with no cited lines, when its skeleton is a tautology. This is Definition 4.8.1 of Section 4.8, applied to skeletons. By Proposition 6.3.1 the line abbreviates a block of primitive steps, pasted in with any flags renamed to names new to the whole derivation, as "The laws as one line" (Section 6.2) allows. As in Section 4.8, use TAUT only for a step a written proof would leave unsaid.

These are TAUT steps:

- from $\neg(\exists x\, P(x) \wedge Q(c))$ and $Q(c)$ to $\neg\exists x\, P(x)$, since $\neg(p \wedge q), q \models \neg p$;
- from $\forall x\,(P(x) \to Q(x)) \to R$ and $\neg R$ to $\neg\forall x\,(P(x) \to Q(x))$: modus tollens on the skeleton, $p \to r, \neg r \models \neg p$.

These are not:

- from $\forall x\,\forall y\, R(x,y)$ to $\forall y\,\forall x\, R(x,y)$: the two prime formulas do not differ by renaming, so the skeletons are p and q, and $p \not\models q$. The step is derivable, by Example (Flagged boxes at work) (c), and that derivation looks inside the quantifiers;
- from $\forall x\, P(x)$ to $P(c)$: the skeletons are p and q; this is $\forall$E;
- $x = x$: the skeleton is one atom, not a tautology; this is =I.

Quick check. Which are TAUT steps? (a) From $\exists x\, P(x) \wedge \forall y\, Q(y)$ to $\forall y\, Q(y) \wedge \exists x\, P(x)$. (b) From $\forall x\,(P(x) \wedge Q(x))$ to $\forall x\, P(x)$. (c) From $\neg\forall x\, P(x)$ to $\exists x\, \neg P(x)$.

Solution.

(a) Yes: $p \wedge q \models q \wedge p$. (b) No: the skeletons are two different atoms; the step is Proposition 6.2.6 (a), which looks inside $\forall x$. (c) No: the skeletons are $\neg p$ and q; the step is QN (a) of Section 6.2, which needs the quantifier rules and RAA.

Warning (Not allowed: TAUT).

- From $\forall x\,(P(x) \wedge Q(x))$ to $\forall x\,Q(x)$. *Looks inside a quantifier; needs a flagged box.*
- From $\forall x\,P(x)$ to $P(f(c))$. *An instance, not an entailment of skeletons; needs* $\forall$E.
- From $P(c)$ to $\exists x\,P(x)$. *Needs* $\exists$I.
- From $s = t$ and $P(s)$ to $P(t)$. *The skeleton is* $p, q \models r$, *which fails; needs* $=$E.
- Citing a line inside a box that has closed. *Not available, whatever the table says (Definition 1.4.2).*

6.3.4 The propositional core

Many first-order derivations have the same shape. Open the boxes the goal asks for, instantiate every $\forall$ line at the names in play, then finish with a step that is propositional. The quantifiers frame the derivation, and its core is a truth-table fact.

Example. A propositional core inside a flagged box

Show that $\forall x\,(P(x) \to (Q(x) \vee R(x)))$, $\forall x\,\neg Q(x) \vdash \forall x\,(P(x) \to R(x))$.

Solution.

1		$\forall x\,(P(x) \to (Q(x) \vee R(x)))$	
2		$\forall x\,\neg Q(x)$	
3	x_0	$P(x_0) \to (Q(x_0) \vee R(x_0))$	$\forall$E 1
4		$\neg Q(x_0)$	$\forall$E 2
5		$P(x_0) \to R(x_0)$	TAUT 3, 4
6		$\forall x\,(P(x) \to R(x))$	$\forall$I 3–5

The skeleton of line 5's step is $p \to (q \vee r), \neg q \models p \to r$. At line 6, x_0 is not in a premise or in line 6, and no assumption is open.

Read over $\mathbb{Z}$, with $P(n)$ as "$n > 1$", $Q(n)$ as "$n^2 = 2$" and $R(n)$ as "$n > 0$". The first premise says "if $n > 1$, then $n^2 = 2$ or $n > 0$", true since $n > 1$ gives $n > 0$; the second says that no integer squares to 2. Both premises are true under this reading, and so is the conclusion, which says that every integer greater than 1 is positive. The quantifier part of the derivation is two instantiations and one generalization, and its core is propositional.

The order of the steps is the point. TAUT could not have been applied to the premises them-

selves: their skeletons are two atoms p and q, and the goal's is a third, r, and $p, q \not\models r$. Only after the flagged box is open and both premises are instantiated at its flag do the lines show the connectives that the table needs. So the move reads: boxes first, instances at the flag second, and the table last.

Quick check. Write out what line 5 abbreviates, with the rules of Chapters 1 and 2.

Solution. Lines 5–13 below replace the one TAUT line:

1		$\forall x\,(P(x) \to (Q(x) \vee R(x)))$	
2		$\forall x\,\neg Q(x)$	
3	x_0	$P(x_0) \to (Q(x_0) \vee R(x_0))$	$\forall$E 1
4		$\neg Q(x_0)$	$\forall$E 2
5		$P(x_0)$	asm, for $\to$I
6		$Q(x_0) \vee R(x_0)$	$\to$E 3, 5
7		$Q(x_0)$	asm, for $\vee$E
8		$\bot$	$\neg$E 7, 4
9		$R(x_0)$	$\bot$E 8
10		$R(x_0)$	asm, for $\vee$E
11		$R(x_0)$	R 10
12		$R(x_0)$	$\vee$E 6, 7–9, 10–11
13		$P(x_0) \to R(x_0)$	$\to$I 5–12
14		$\forall x\,(P(x) \to R(x))$	$\forall$I 3–13

That is nine lines in place of one. A written proof would say "so $R(x_0)$, since $\neg Q(x_0)$".

6.3.5 Four kinds of task, with quantifiers

Section 3.2 sorted propositional problems into four kinds, and the kind decides how the search begins. The same four come back. A **theorem** $\vdash \varphi$ has no premises; an **equivalence** $\varphi \dashv\vdash \psi$ needs two derivations; **inconsistency** of a list is a derivation of $\bot$ from it (Section 3.2's definition reads the same for first-order formulas, now with all the rules); and **derive or break** withholds the verdict. A claim is broken by a counter-reading, which makes every member of its counterexample set true (Section 7.2).

Example. One pair of predicates, four tasks

(a) *Theorem.* Show that $\vdash \forall x\,P(x) \to \exists x\,(P(x) \vee Q(x))$.

(b) *Equivalence.* Show that $\forall x\,(P(x) \to \neg Q(x)) \dashv\vdash \forall x\,(Q(x) \to \neg P(x))$.

(c) *Inconsistency.* Show that the list $\forall x\,(P(x) \to Q(x))$, $\exists x\,(P(x) \wedge \neg Q(x))$ is inconsistent.

(d) *Derive or break.* Decide $\forall x\,(P(x) \to Q(x)) \vdash \forall x\,P(x) \to \forall x\,Q(x)$ and its converse.

Solution. (a) With no premises the search starts from the goal alone: an $\to$I box, then the $\exists$ goal waits for a line that shows an instance.

1		$\forall x\,P(x)$	asm, for $\to$I
2		$P(x)$	$\forall$E 1
3		$P(x) \lor Q(x)$	$\lor$I 2
4		$\exists x\,(P(x) \lor Q(x))$	$\exists$I 3
5	$\forall x\,P(x) \to \exists x\,(P(x) \lor Q(x))$		$\to$I 1–4

Line 2 instantiates at the variable x itself, and line 4 uses $\exists$I at x, as in Proposition 5.5.1. Any term would have done as the witness, since line 1 holds at every term; the variable x is simply the term at hand, and no constant has to be invented.

(b) The E form of Section 6.2 says the same with its two predicates exchanged: "no P is Q" is "no Q is P". An equivalence is two searches, and here the second is the first with the letters exchanged, so one search does the work of both. Each direction opens a flagged box, instantiates at its flag, takes one TAUT step and generalizes:

1		$\forall x\,(P(x) \to \neg Q(x))$	
2	x_0	$P(x_0) \to \neg Q(x_0)$	$\forall$E 1
3		$Q(x_0) \to \neg P(x_0)$	TAUT 2
4		$\forall x\,(Q(x) \to \neg P(x))$	$\forall$I 2–3

1		$\forall x\,(Q(x) \to \neg P(x))$	
2	x_0	$Q(x_0) \to \neg P(x_0)$	$\forall$E 1
3		$P(x_0) \to \neg Q(x_0)$	TAUT 2
4		$\forall x\,(P(x) \to \neg Q(x))$	$\forall$I 2–3

The skeletons are $p \to \neg q \models q \to \neg p$ and the same with p and q exchanged.

(c) An A form and its O form (Section 6.2). The goal is $\bot$, so the clash step of the method applies, but no pair χ, $\neg\chi$ is in sight among the premises: the clash lives at a witness. The $\exists$ premise opens its box at once, the $\forall$ premise is instantiated at the witness, and the clash is propositional:

1		$\forall x\,(P(x) \to Q(x))$	
2		$\exists x\,(P(x) \land \neg Q(x))$	
3	x_0	$P(x_0) \land \neg Q(x_0)$	asm, for $\exists$E
4		$P(x_0) \to Q(x_0)$	$\forall$E 1
5		$\bot$	TAUT 3, 4
6		$\bot$	$\exists$E 2, 3–5

Line 5's skeleton is $p \land \neg q, p \to q \models \bot$.

(d) With the verdict withheld, spend a minute on each side before committing. For the first claim, a reading that makes the premise true and the conclusion false would need every P to be a Q, every element to be a P, and some element not a Q, which cannot happen; so try to derive it. For the converse, a false hypothesis $\forall x\,P(x)$ makes the premise true cheaply,

which suggests a break.

The first claim derives. Lines 3–6 are Example (Flagged boxes at work) (b), run inside the box of line 2:

1	$\forall x\,(P(x) \to Q(x))$		
2	$\quad \forall x\,P(x)$		asm, for $\to$I
3	$\quad x_0 \mid P(x_0)$		$\forall$E 2
4	$\quad\quad P(x_0) \to Q(x_0)$		$\forall$E 1
5	$\quad\quad Q(x_0)$		$\to$E 4, 3
6	$\quad \forall x\,Q(x)$		$\forall$I 3–5
7	$\forall x\,P(x) \to \forall x\,Q(x)$		$\to$I 2–6

At line 6, x_0 is not in the premise, in line 6, or in the open assumption on line 2.

The converse breaks. Over $\mathbb{Z}$, read P as "even" and Q as "positive". "Every integer is even" is false, so the premise "if every integer is even, then every integer is positive" is true; the conclusion "every even integer is positive" is false at 0.

Test the subgoal by a counter-reading. Section 4.7 tested a propositional subgoal by a refuting row before chasing it. The first-order form of the move is the same. When a search produces a subgoal, look for a reading that makes the available lines true and the subgoal false before you chase it. If you find one, the subgoal is the wrong one: Section 7.2 states that such a reading rules out every derivation of it from those lines. Return to the step that produced it, and choose differently there: another term for a $\forall E$, a box opened earlier, or a different case split. A reading takes a minute to find, and it can save the search a long chase.

Quick check. A search for the limit derivation of Section 6.1 instantiates the premise at ε_0 instead of $\varepsilon_0/2$, and reaches the subgoal $|2a_{n_0} - 2L| < \varepsilon_0$ with $|a_{n_0} - L| < \varepsilon_0$ available. Find a sequence and numbers that make every available line true and the subgoal false.

Solution. Take $\varepsilon_0 = 1$; $a_n = L$ for every n except $a_{n_0} = L + 0.75$; and $N_0 = n_0$. The premise holds, since the sequence is eventually constant at L. Every term is within $0.75 < 1$ of L, so the instance of the premise at ε_0 and the box line $\forall n \in \mathbb{N}\ (n \geq N_0 \to |a_n - L| < \varepsilon_0)$ hold; $n_0 \geq N_0$ holds; and $|a_{n_0} - L| = 0.75 < 1$. But $|2a_{n_0} - 2L| = 1.5 \geq 1$. The subgoal is the wrong one; go back to the choice of term.

Quick check. Which task is $\exists x\,(P(x) \wedge \neg P(x)) \vdash \bot$, and how long is its derivation with TAUT?

Solution. Inconsistency. Three lines below the premise:

1	$\exists x\,(P(x) \wedge \neg P(x))$	
2	x_0 | $P(x_0) \wedge \neg P(x_0)$	asm, for $\exists$E
3	| $\bot$	TAUT 2
4	$\bot$	$\exists$E 1, 2–3

6.3.6 A library of first-order sequents

Here is what the book has proved about quantifiers, grouped by use; a search that needs one of these can copy its derivation, with formulas in place of letters, rather than rediscover it.

Use	Sequent	Where
exchange two like quantifiers	$\forall x\,\forall y\,\varphi \dashv\vdash \forall y\,\forall x\,\varphi$	Example (Flagged boxes at work) (c)
rename a bound variable	$\forall y\,\varphi \dashv\vdash \forall z\,\varphi[z/y]$, and the same for $\exists$	Section 5.4 (the fold after Example (Flagged boxes at work)); Section 5.6
every gives some	$\forall x\,\varphi \vdash \exists x\,\varphi$	Proposition 5.5.1
some-all gives all-some	$\exists x\,\forall y\,R(x,y) \vdash \forall y\,\exists x\,R(x,y)$	Proposition 5.6.2
distribute	$\forall$ over $\wedge$ and $\exists$ over $\vee$ both ways; $\forall$ over $\vee$ and $\exists$ over $\wedge$ one way	Proposition 6.2.6 (the instances for P, Q first in Proposition 5.4.1 and Proposition 5.6.1)
negate	"not every" and "some not"; "none" and "every not"; a negated implication; bounded forms	Proposition 6.2.1, Proposition 6.2.2, Proposition 6.2.3
move across a connective	the passage laws	Proposition 6.2.7
replace a part	a provable equivalence, inside any formula	Proposition 6.2.4
equality	symmetric and transitive; equals share every property	Proposition 5.7.1, Proposition 5.7.2
bounds in one step	the bounded rules	Proposition 6.1.1
propositional steps	TAUT on the skeleton	Proposition 6.3.1

The laws of Section 6.2 may already be used as one line, by "The laws as one line"; the other rows are patterns to copy. Either way, a row is used by matching its shape and then checking its conditions: the passage laws need the moved variable not free in the other side, a renaming needs a letter new to the formula, and an instance of a row's letters must be free for the variables it replaces.

Remark (Looking ahead). Every claim this section broke, it broke by a counter-reading. Chapter 7 gives counter-readings an exact form, the counter-model, and states that a counter-model rules out every derivation.

6.3.7 Exercises

6.3.7.1 A. Check your understanding

Exercise. A1

(a) What is the propositional skeleton of $\neg\exists x\, P(x) \lor (Q(c) \land \exists y\, P(y))$? Is it a tautology?

(b) What is the propositional skeleton of $\exists x\, (P(x) \land Q(x)) \land \neg\exists x\, (Q(x) \land P(x))$? Can a TAUT step derive $\bot$ from this formula?

(c) Which are TAUT steps? (i) From $\forall x\, P(x)$ and $\forall x\, P(x) \to \exists y\, Q(y)$ to $\exists y\, Q(y)$. (ii) From $\forall x\, (P(x) \to Q(x))$ and $P(c)$ to $Q(c)$. (iii) From $\neg\neg\exists x\, P(x)$ to $\exists x\, P(x)$.

(d) In Example (A search that waits for its witnesses), why is the box on premise 3 opened before premise 1 is used?

(e) A counter-reading makes the premises of a claim true and its conclusion false. What may you say now, and what will Section 7.2 add?

Solution. *(a)* $\neg p \lor (q \land p)$: $\exists x\, P(x)$ and $\exists y\, P(y)$ differ only by renaming, so they share p. It is not a tautology: p true and q false make it false.

(b) $p \land \neg q$. The two prime formulas are not renamings of each other ($P(x) \land Q(x)$ against $Q(x) \land P(x)$), so they get different atoms, and $p \land \neg q$ is true in a row: no. $\bot$ is derivable all the same, by an $\exists$E box on the first conjunct, a TAUT step inside it to $Q(x_0) \land P(x_0)$, $\exists$I, and a clash with the second conjunct.

(c) (i) Yes: $p, p \to q \models q$. (ii) No: $p, q \not\models r$; it is $\forall$E at c, then $\to$E. (iii) Yes: $\neg\neg p \models p$.

(d) Premise 1 is a $\forall$ whose useful term is not yet known, and the box supplies it, x_0. These are the moves "name the witness early" and "instantiate at the term you need", in that order.

(e) Now: the claim breaks, and a search for it is abandoned. Section 7.2 states that such a claim has no derivation.

6.3.7.2 B. Practice

Exercise. B1: Derive or break, twelve claims

Derive each claim, with TAUT allowed, or break it by a reading over $\mathbb{Z}$ that makes every member of its counterexample set true. The groups say which move to try first; they do not say which claims hold. Claims 1–3: name the witness early. Claims 4 and 5: instantiate at the term you need. Claims 6–8: the propositional core. Claims 9–12: move a quantifier, or derive $\bot$.

1. $\exists x\, (P(x) \land Q(x)),\ \forall x\, (P(x) \to (Q(x) \to R(f(x)))) \vdash \exists y\, R(y)$.
2. $\exists x\, R(x, x) \vdash \forall x\, \exists y\, R(x, y)$.
3. $\exists x\, P(x),\ \forall x\, (P(x) \to \forall y\, R(x, y)) \vdash \forall y\, \exists x\, R(x, y)$.
4. $\forall x\, \forall y\, (R(x, y) \to R(y, x)),\ \forall x\, R(x, f(x)) \vdash \forall x\, R(f(x), x)$.
5. $\forall x\, R(x, f(x)) \vdash \forall x\, R(f(x), x)$.
6. $\forall x\, (P(x) \to Q(x)),\ \forall x\, (Q(x) \to R(x)) \vdash \forall x\, (\neg R(x) \to \neg P(x))$.
7. $\forall x\, (P(x) \lor Q(x)),\ \forall x\, \neg P(x) \vdash \forall x\, Q(x)$.

8. $\forall x\,(P(x) \lor Q(x)),\ \neg\forall x\,P(x) \vdash \forall x\,Q(x)$.
9. $\forall x\,(P(x) \to Q(x)) \vdash \exists x\,P(x) \to \exists x\,Q(x)$.
10. $\forall x\,(P(x) \to Q(x)) \vdash \exists x\,P(x) \to \forall x\,Q(x)$.
11. $\forall x\,(P(x) \leftrightarrow \neg P(f(x))),\ \forall x\,P(x) \vdash \bot$.
12. $\forall x\,\exists y\,R(x,y),\ \forall x\,\forall y\,(R(x,y) \to \neg R(y,x)) \vdash \bot$.

Solution. 1. Derive. The box of the witness opens first; inside it, premise 2 is instantiated at x_0, and the postponed goal takes the compound witness $f(x_0)$.

1		$\exists x\,(P(x) \land Q(x))$	
2		$\forall x\,(P(x) \to (Q(x) \to R(f(x))))$	
3	x_0	$P(x_0) \land Q(x_0)$	asm, for $\exists$E
4		$P(x_0) \to (Q(x_0) \to R(f(x_0)))$	$\forall$E 2
5		$R(f(x_0))$	TAUT 3, 4
6		$\exists y\,R(y)$	$\exists$I 5
7		$\exists y\,R(y)$	$\exists$E 1, 3–6

Line 5's skeleton is $p \land q, p \to (q \to r) \models r$. At line 7, x_0 is not in line 7, in line 1, or in a premise.

2. Breaks. Read $R(x,y)$ as "$x \cdot y = 1$". The premise is true, since $1 \cdot 1 = 1$; the conclusion fails at $x = 0$, since $0 \cdot y = 0$ for every y.

3. Derive.

1		$\exists x\,P(x)$	
2		$\forall x\,(P(x) \to \forall y\,R(x,y))$	
3	x_0	$P(x_0)$	asm, for $\exists$E
4		$P(x_0) \to \forall y\,R(x_0,y)$	$\forall$E 2
5		$\forall y\,R(x_0,y)$	$\to$E 4, 3
6	y_0	$R(x_0,y_0)$	$\forall$E 5
7		$\exists x\,R(x,y_0)$	$\exists$I 6
8		$\forall y\,\exists x\,R(x,y)$	$\forall$I 6–7
9		$\forall y\,\exists x\,R(x,y)$	$\exists$E 1, 3–8

At line 8, y_0 is not in a premise, in the open assumption on line 3, or in line 8. At line 9, x_0 is not in line 9, in line 1, or in a premise.

4. Derive. The term premise 1 needs is $f(x_0)$, and premise 2 names it.

1		$\forall x\, \forall y\, (R(x,y) \to R(y,x))$	
2		$\forall x\, R(x, f(x))$	
3	x_0	$R(x_0, f(x_0))$	$\forall$E 2
4		$\forall y\, (R(x_0,y) \to R(y,x_0))$	$\forall$E 1
5		$R(x_0, f(x_0)) \to R(f(x_0), x_0)$	$\forall$E 4
6		$R(f(x_0), x_0)$	$\to$E 5, 3
7		$\forall x\, R(f(x), x)$	$\forall$I 3–6

Line 5 instantiates at $f(x_0)$, which is free for y. At line 7, x_0 is not in a premise or in line 7, and no assumption is open.

5. Breaks. Read $f(x)$ as $x + 1$ and R as $<$. Every $x < x + 1$, so the premise is true; no $x + 1 < x$, so the conclusion is false.

6. Derive.

1		$\forall x\, (P(x) \to Q(x))$	
2		$\forall x\, (Q(x) \to R(x))$	
3	x_0	$P(x_0) \to Q(x_0)$	$\forall$E 1
4		$Q(x_0) \to R(x_0)$	$\forall$E 2
5		$\neg R(x_0) \to \neg P(x_0)$	TAUT 3, 4
6		$\forall x\, (\neg R(x) \to \neg P(x))$	$\forall$I 3–5

Line 5's skeleton is $p \to q, q \to r \models \neg r \to \neg p$.

7. Derive.

1		$\forall x\, (P(x) \lor Q(x))$	
2		$\forall x\, \neg P(x)$	
3	x_0	$P(x_0) \lor Q(x_0)$	$\forall$E 1
4		$\neg P(x_0)$	$\forall$E 2
5		$Q(x_0)$	TAUT 3, 4
6		$\forall x\, Q(x)$	$\forall$I 3–5

Line 5's skeleton is $p \lor q, \neg p \models q$.

8. Breaks. Read P as "even" and Q as "odd". Every integer is even or odd; not every integer is even; and 0 is not odd, so "every integer is odd" is false.

9. Derive.

1	$\forall x\,(P(x) \to Q(x))$	
2	$\quad \exists x\,P(x)$	asm, for $\to$I
3	$\quad x_0 \mid P(x_0)$	asm, for $\exists$E
4	$\quad\quad P(x_0) \to Q(x_0)$	$\forall$E 1
5	$\quad\quad Q(x_0)$	$\to$E 4, 3
6	$\quad\quad \exists x\,Q(x)$	$\exists$I 5
7	$\quad \exists x\,Q(x)$	$\exists$E 2, 3–6
8	$\exists x\,P(x) \to \exists x\,Q(x)$	$\to$I 2–7

At line 7, x_0 is not in line 7, in line 2, in the premise, or in an open assumption (line 2's box is still open, and line 2 does not mention x_0).

10. Breaks. Read $P(x)$ as "$x = 0$" and $Q(x)$ as "$x \cdot x = x$". The premise holds, since the only x with $x = 0$ has $0 \cdot 0 = 0$. In the conclusion, $\exists x\,(x = 0)$ is true and $\forall x\,(x \cdot x = x)$ is false, since $2 \cdot 2 = 4$; so the implication is false.

11. Derive. Premise 2 is instantiated twice, at y and at $f(y)$, the two terms that premise 1 at y mentions.

1	$\forall x\,(P(x) \leftrightarrow \neg P(f(x)))$	
2	$\forall x\,P(x)$	
3	$P(y) \leftrightarrow \neg P(f(y))$	$\forall$E 1
4	$P(y)$	$\forall$E 2
5	$P(f(y))$	$\forall$E 2
6	$\bot$	TAUT 3, 4, 5

Line 6's skeleton is $p \leftrightarrow \neg q, p, q \models \bot$: no row makes all three true.

12. Breaks. Read R as $<$. Every integer has a larger one, and $x < y$ excludes $y < x$. Both premises are true, and $\neg\bot$ is true, so the whole counterexample set is true: the list is not shown inconsistent.

Seven claims derive and five break. Four breaks sit beside a derivable look-alike: 5 beside 4, 8 beside 7, 10 beside 9, and 12 beside 11. Claim 2 weakens the premise of the derivable $\forall x\,R(x,x) \vdash \forall x\,\exists y\,R(x,y)$ of Example (Derivations with existential introduction) (c). Every break is a counter-reading (Section 7.2).

Marking scheme. (10 points) - 4: the twelve verdicts, $\frac{1}{3}$ point for each correct one, rounded down. - 3: the seven derivations, each with every closing side condition checked and every TAUT line's skeleton named. - 3: the five readings, each with every premise and the conclusion evaluated. Fatal errors (0 for the claim): a TAUT line that looks inside a quantifier; a "break" with a premise false.

Exercise. B2: Two witnesses in a row

The premises are 1 $\forall x\,(P(x) \to \exists y\,R(x,y))$, 2 $\forall x\,\forall y\,(R(x,y) \to P(y))$ and 3 $P(c)$; the goal is $\exists y\,\exists z\,(R(c,y) \wedge R(y,z))$. Here is a partial derivation:

1		$\forall x\,(P(x) \to \exists y\,R(x,y))$	
2		$\forall x\,\forall y\,(R(x,y) \to P(y))$	
3		$P(c)$	
		⋮	
6	y_0	$R(c,y_0)$	asm, for $\exists$E
		⋮	
16		$\exists y\,\exists z\,(R(c,y) \wedge R(y,z))$	?
17		$\exists y\,\exists z\,(R(c,y) \wedge R(y,z))$	?

Name the move, then complete it.

Solution. The move is "name the witness early", twice, with each $\forall$ line instantiated only at the terms in hand: c (premise 3) first, then the witness y_0.

1		$\forall x\,(P(x) \to \exists y\,R(x,y))$	
2		$\forall x\,\forall y\,(R(x,y) \to P(y))$	
3		$P(c)$	
4		$P(c) \to \exists y\,R(c,y)$	$\forall$E 1
5		$\exists y\,R(c,y)$	$\to$E 4, 3
6	y_0	$R(c,y_0)$	asm, for $\exists$E
7		$\forall y\,(R(c,y) \to P(y))$	$\forall$E 2
8		$R(c,y_0) \to P(y_0)$	$\forall$E 7
9		$P(y_0)$	$\to$E 8, 6
10		$P(y_0) \to \exists y\,R(y_0,y)$	$\forall$E 1
11		$\exists y\,R(y_0,y)$	$\to$E 10, 9
12		z_0 | $R(y_0,z_0)$	asm, for $\exists$E
13		$R(c,y_0) \wedge R(y_0,z_0)$	$\wedge$I 6, 12
14		$\exists z\,(R(c,y_0) \wedge R(y_0,z))$	$\exists$I 13
15		$\exists y\,\exists z\,(R(c,y) \wedge R(y,z))$	$\exists$I 14
16		$\exists y\,\exists z\,(R(c,y) \wedge R(y,z))$	$\exists$E 11, 12–15
17		$\exists y\,\exists z\,(R(c,y) \wedge R(y,z))$	$\exists$E 5, 6–16

Lines 4 and 7 instantiate at c, and lines 8 and 10 at y_0; line 15 is $\exists$I at y_0, which is free for y. At line 16, z_0 is not in line 16, in line 11, in the premises, or in the open assumption on line 6. At line 17, y_0 is not in line 17, in line 5, or in the premises.

Marking scheme. (10 points) - 2: names the move. - 3: opens the two boxes as soon as their $\exists$ lines exist. - 3: the $\forall$E lines at c, c, y_0 and y_0, the terms read off the lines in hand. - 2: the two side conditions. Fatal error (0 for the step): $\exists$I on line 15 before line 14, so that the witness z_0 appears in the box's last line.

Exercise. B3: How much TAUT saves

(a) Derive $\forall x\,(P(x) \to Q(x)),\ \forall x\,(Q(x) \to \neg P(x)) \vdash \forall x\,\neg P(x)$ twice: once without TAUT, once with exactly one TAUT line. Count the lines.

(b) Is the step from $\forall x\,(P(x) \to Q(x))$ to $\forall x\,(\neg Q(x) \to \neg P(x))$ a TAUT step?

Solution. (a) With TAUT, six lines (left); without it, ten (right):

1		$\forall x\,(P(x) \to Q(x))$	
2		$\forall x\,(Q(x) \to \neg P(x))$	
3	x_0	$P(x_0) \to Q(x_0)$	$\forall$E 1
4		$Q(x_0) \to \neg P(x_0)$	$\forall$E 2
5		$\neg P(x_0)$	TAUT 3, 4
6		$\forall x\,\neg P(x)$	$\forall$I 3–5

1		$\forall x\,(P(x) \to Q(x))$	
2		$\forall x\,(Q(x) \to \neg P(x))$	
3	x_0	$P(x_0) \to Q(x_0)$	$\forall$E 1
4		$Q(x_0) \to \neg P(x_0)$	$\forall$E 2
5		$P(x_0)$	asm, for $\neg$I
6		$Q(x_0)$	$\to$E 3, 5
7		$\neg P(x_0)$	$\to$E 4, 6
8		$\bot$	$\neg$E 5, 7
9		$\neg P(x_0)$	$\neg$I 5–8
10		$\forall x\,\neg P(x)$	$\forall$I 3–9

Line 5 on the left has the skeleton $p \to q, q \to \neg p \models \neg p$.

(b) No: the two formulas are prime and not renamings of each other, so their skeletons are two different atoms. The step is derivable by instantiating, one TAUT line ($p \to q \models \neg q \to \neg p$) and generalizing, the pattern of Example (One pair of predicates, four tasks) (b).

Marking scheme. (10 points) - 3: the TAUT derivation, with the skeleton named. - 4: the derivation without TAUT. - 1: the count. - 2: (b), with the reason. Fatal error (0 for the step): a TAUT line citing the $\forall$ lines directly to reach $\neg P(x_0)$.

6.3.7.3 C. Going deeper

Exercise. C1: Some index decides, used later

Let $D(n)$ say "$a_n > 0$" for a sequence (a_n), over a domain of indices.

(a) Derive $\vdash \exists n\,(D(n) \to \forall m\, D(m))$, with LEM on $\forall m\, D(m)$.

(b) Derive it again in two lines, from one TAUT line and one law of Section 6.2.

(c) Say in two sentences which index the sentence promises, and why the derivations of (a) and (b) name none.

Solution (Hint 1). For (a): in the case $\neg\forall m\, D(m)$, which law of Section 6.2 turns the negated $\forall$ into an $\exists$?

Solution (Hint 2). For (b): the formula $\forall n\, D(n) \to \forall m\, D(m)$ has the skeleton $p \to p$. Which passage law moves the $\forall n$ out of the hypothesis of an implication, turning it into $\exists n$?

Solution. (a)

1	$\forall m\, D(m) \lor \neg\forall m\, D(m)$	LEM
2	$\quad \forall m\, D(m)$	asm, for $\lor$E
3	$\quad\quad D(n)$	asm, for $\to$I
4	$\quad\quad \forall m\, D(m)$	R 2
5	$\quad D(n) \to \forall m\, D(m)$	$\to$I 3–4
6	$\quad \exists n\,(D(n) \to \forall m\, D(m))$	$\exists$I 5
7	$\quad \neg\forall m\, D(m)$	asm, for $\lor$E
8	$\quad \exists m\, \neg D(m)$	QN (a) 7
9	$\quad m_0 \mid \neg D(m_0)$	asm, for $\exists$E
10	$\quad\quad D(m_0) \to \forall m\, D(m)$	VI 9
11	$\quad\quad \exists n\,(D(n) \to \forall m\, D(m))$	$\exists$I 10
12	$\quad \exists n\,(D(n) \to \forall m\, D(m))$	$\exists$E 8, 9–11
13	$\exists n\,(D(n) \to \forall m\, D(m))$	$\lor$E 1, 2–6, 7–12

Line 6 is $\exists$I at the variable n itself, as in the proof of passage law (d); as there, the step is sound because every domain is non-empty (Section 7.1). Line 11 is $\exists$I at m_0. At line 12, m_0 is not in line 12, in line 8, or in the open assumption on line 7.

(b)

1	$\forall n\, D(n) \to \forall m\, D(m)$	TAUT
2	$\exists n\, (D(n) \to \forall m\, D(m))$	passage (d) 1

The two prime formulas of line 1 differ only by renaming, so the skeleton is $p \to p$, a tautology. Line 2 is Proposition 6.2.7 (d), the direction $\dashv$, used as one line ("The laws as one line"), with φ the formula $D(n)$ and ψ the formula $\forall m\, D(m)$, in which n is not free.

(c) If every term is positive, any index will do; if not, an index with $a_n \leq 0$ does, since then the implication is vacuously true. The derivation splits on which case holds without deciding it (LEM in (a), the classical direction of (d) in (b)), so it never names the index.

Marking scheme. (10 points) - 4: (a), with the LEM split, both cases closed, with the side condition of every line that closes a flagged box checked. - 3: (b), with the TAUT line justified by renaming, and law (d) with its condition checked. - 3: (c). Fatal errors (0 for the step): giving $\forall n\, D(n)$ and $\forall m\, D(m)$ different atoms and then claiming the TAUT line anyway; $\exists$I at m_0 after its box has closed.

Exercise. C2: Every M is P, every M is S, so some S is P?

(a) Break $\forall x\, (M(x) \to P(x)),\ \forall x\, (M(x) \to S(x)) \vdash \exists x\, (S(x) \wedge P(x))$ by a reading over $\mathbb{N}$.

(b) Add the premise $\exists x\, M(x)$ and derive the conclusion.

(c) Which line of (b) needs the new premise, and what does Section 5.1's existential import have to do with it?

Solution. (a) Breaks. Read $M(n)$ as "n is an even prime greater than 2", which is true of no n; $P(n)$ as "n is odd"; and $S(n)$ as "n is even". Both premises are vacuously true, since their hypothesis $M(n)$ is false for every n, and "some even number is odd" is false (Section 7.2).

(b)

1		$\forall x\, (M(x) \to P(x))$	
2		$\forall x\, (M(x) \to S(x))$	
3		$\exists x\, M(x)$	
4	x_0	$M(x_0)$	asm, for $\exists$E
5		$M(x_0) \to P(x_0)$	$\forall$E 1
6		$M(x_0) \to S(x_0)$	$\forall$E 2
7		$S(x_0) \wedge P(x_0)$	TAUT 4, 5, 6
8		$\exists x\, (S(x) \wedge P(x))$	$\exists$I 7
9		$\exists x\, (S(x) \wedge P(x))$	$\exists$E 3, 4–8

Line 7's skeleton is $m, m \to p, m \to s \models s \wedge p$. At line 9, x_0 is not in line 9, in line 3, or in a premise.

(c) Line 7: its TAUT needs $M(x_0)$, line 4, which only the box of premise 3 supplies. Without

it, ∃I could still be applied at some term, but no line would say that term is S and P. A "for every" claims no instance, so the old reading of "every M is P" as also saying that there are Ms is not the book's reading. With that reading the inference was valid; without it, the premise must be added.

Marking scheme. (10 points) - 3: (a), with all three truth values justified. - 4: (b). - 3: (c). Fatal error (0 for the step): a "break" whose M is true of some number.

6.4 Uniformly Continuous Implies Continuous, Rule by Rule

Chapter 5 opened with a question. Continuity and uniform continuity differ only in where one quantifier stands; why does one imply the other in a few lines, while the converse is false? Section 5.1 gave the informal answer: a variable may depend on the variables quantified to its left, and only on them. Exercise (C1: Where the $\forall x$ stands) of Section 6.1 asked for both translations into formulas, and Section 6.2 negated uniform continuity (Example (Three negations from analysis) (c)). This section answers the question with the rules. It derives "uniformly continuous implies continuous" in full, line by line, with the bounded rules of Proposition 6.1.1. It then reads off the derivation the one thing this section adds: **which boxes are open where a constant is chosen**, and what that says about what the constant may depend on. The same reading locates the exact lines at which the converse breaks down, and the function x^2 shows that the converse is not merely unproved but false.

6.4.1 Two definitions that differ in one place

Let A be a set of real numbers and f a function defined on A. We work over the domain $\mathbb{R}$, with f a one-place function symbol, $\in A$ a one-place predicate symbol (Section 5.2), and $|\cdot|$ the absolute value. Here the two definitions are only **quantifier patterns**, and nothing is used about f except that $f(x)$ is a term. (A function symbol has a value at every real number; the values outside A are never used, because f is only applied to variables bounded by $\in A$, as with sequences in Section 5.1.)

f is **uniformly continuous on** A (UC) when

$$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A\ \forall y \in A\ (|x - y| < \delta \to |f(x) - f(y)| < \varepsilon).$$

f is **continuous at every point of** A (C) when

$$\forall x \in A\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall y \in A\ (|x - y| < \delta \to |f(x) - f(y)| < \varepsilon).$$

The two formulas have the same matrix and the same four bounded quantifiers. Only $\forall x \in A$ has moved: in C it stands **before** $\exists \delta > 0$, in UC **after** it. In the game of Section 5.1, the defender chooses δ in C after seeing x and ε, and in UC after seeing only ε. So in UC a single δ must serve every point of A at once: that is what **uniformly** means.

Quick check. Swap the first two quantifiers of UC, to get $\exists \delta > 0\ \forall \varepsilon > 0\ \forall x \in A\ \forall y \in A\ (|x - y| < \delta \to |f(x) - f(y)| < \varepsilon)$. What does the new formula say about two points of A closer than δ? Does $f(x) = x$ on $\mathbb{R}$ satisfy it?

Solution. Now δ is chosen before ε, so it must work for every $\varepsilon > 0$ at once. For two points $x, y \in A$ with $|x-y| < \delta$, this says $|f(x)-f(y)| < \varepsilon$ for every $\varepsilon > 0$, and by Proposition 6.2.5, $f(x) = f(y)$. So the formula says: for some $\delta > 0$, any two points of A closer than δ have the **same** value. The function $f(x) = x$ on $\mathbb{R}$ fails it: whatever $\delta > 0$ is proposed, the points 0 and $\delta/2$ are closer than δ and have different values. Yet $f(x) = x$ is uniformly continuous, with $\delta = \varepsilon$. Swapping $\forall\varepsilon$ and $\exists\delta$ asks far too much.

6.4.2 The skeleton first

Hold ε fixed, and write $K(x, \delta)$ for "δ works at x", the formula $\forall y \in A\ (|x - y| < \delta \to |f(x) - f(y)| < \varepsilon)$. Drop the bounds on δ and x for a moment. Then UC at this ε says $\exists\delta\,\forall x\,K(x, \delta)$, one δ for every point, and C at this ε says $\forall x\,\exists\delta\,K(x, \delta)$, a δ for each point. The step from the first to the second is Proposition 5.6.2 with other letters:

1	$\exists\delta\,\forall x\,K(x,\delta)$			
2	x_0	δ_0	$\forall x\,K(x,\delta_0)$	asm, for $\exists$E
3			$K(x_0,\delta_0)$	$\forall$E 2
4			$\exists\delta\,K(x_0,\delta)$	$\exists$I 3
5		$\exists\delta\,K(x_0,\delta)$		$\exists$E 1, 2–4
6	$\forall x\,\exists\delta\,K(x,\delta)$			$\forall$I 2–5

Fix a point, obtain the one δ that works everywhere, choose it for this point. The converse, from $\forall x\,\exists\delta\,K$ to $\exists\delta\,\forall x\,K$, is the swap whose search stalled in Section 5.6. The rest of this section puts ε and the bounds back and writes K out. That costs five lines: the premise UC itself (line 1), the assumptions $x_0 \in A$, $\varepsilon_0 > 0$ and $\delta_0 > 0$ (lines 2, 3 and 5), and the line that closes the box of ε_0 (line 10). The skeleton's line 1 becomes line 4 of the derivation below (UC at ε_0), its lines 2–5 become lines 6–9, and its line 6 becomes line 11, which closes the box of x_0.

Quick check. Why may line 5 of the skeleton close the box of δ_0?

Solution. δ_0 does not appear in line 5, $\exists\delta\,K(x_0, \delta)$, nor in the existential of line 1, which is also the only premise, and the one box still open, that of x_0, has no assumption. Every clause of $\exists$E holds.

6.4.3 The derivation

Example. Uniformly continuous implies continuous

Derive C from UC, with the bounded rules of Proposition 6.1.1.

Idea. Run the search method of Section 6.3. The goal C begins $\forall x \in A$, then $\forall\varepsilon > 0$: the **goal-directed** steps open two bounded $\forall$I boxes, "let $x_0 \in A$" and "let $\varepsilon_0 > 0$", and write

the last two lines. The new goal begins $\exists\delta > 0$, and no witness is in sight, so it is **postponed** (Section 5.5). Turn to the premise UC. It begins $\forall\varepsilon > 0$, and the term the goal mentions is ε_0, whose bound $\varepsilon_0 > 0$ is on a line: $\forall$E at ε_0 (Section 5.3). That gives an $\exists\delta > 0$ line, and its $\exists$E box is opened **at once**, with a witness δ_0 (Section 5.6), aiming for the postponed goal as its last line. Inside, the premise has become "δ_0 works for all $x, y \in A$"; instantiate it at the point we have, x_0. Now the postponed goal has its witness: δ_0 itself.

Before reading on: which lines does the line introducing δ_0 rest on, and does any of them mention x_0?

Solution.

1		$\forall\varepsilon > 0\ \exists\delta > 0\ \forall x \in A\ \forall y \in A\ (\lvert x - y\rvert < \delta \to \lvert f(x) - f(y)\rvert < \varepsilon)$	
2	x_0	$x_0 \in A$	asm, for $\forall$I
3	ε_0	$\varepsilon_0 > 0$	asm, for $\forall$I
4		$\exists\delta > 0\ \forall x \in A\ \forall y \in A\ (\lvert x - y\rvert < \delta \to \lvert f(x) - f(y)\rvert < \varepsilon_0)$	$\forall$E 1, 3
5	δ_0	$\delta_0 > 0$	asm, for $\exists$E
6		$\forall x \in A\ \forall y \in A\ (\lvert x - y\rvert < \delta_0 \to \lvert f(x) - f(y)\rvert < \varepsilon_0)$	asm, for $\exists$E
7		$\forall y \in A\ (\lvert x_0 - y\rvert < \delta_0 \to \lvert f(x_0) - f(y)\rvert < \varepsilon_0)$	$\forall$E 6, 2
8		$\exists\delta > 0\ \forall y \in A\ (\lvert x_0 - y\rvert < \delta \to \lvert f(x_0) - f(y)\rvert < \varepsilon_0)$	$\exists$I 5, 7
9		$\exists\delta > 0\ \forall y \in A\ (\lvert x_0 - y\rvert < \delta \to \lvert f(x_0) - f(y)\rvert < \varepsilon_0)$	$\exists$E 4, 5–8
10		$\forall\varepsilon > 0\ \exists\delta > 0\ \forall y \in A\ (\lvert x_0 - y\rvert < \delta \to \lvert f(x_0) - f(y)\rvert < \varepsilon)$	$\forall$I 3–9
11		$\forall x \in A\ \forall\varepsilon > 0\ \exists\delta > 0\ \forall y \in A\ (\lvert x - y\rvert < \delta \to \lvert f(x) - f(y)\rvert < \varepsilon)$	$\forall$I 2–10

Every line is tagged with its rule; lines 4, 7 and 8 are the one-line bounded rules, and lines 9, 10 and 11 close bounded boxes. The terms and the side conditions:

- Line 4: $\forall$E at ε_0, with its bound on line 3. The term ε_0 is a variable that does not appear in line 1, so it is free for ε.
- Line 7: $\forall$E at x_0, with its bound on line 2, which is available because the box of line 2 is still open.
- Line 8: $\exists$I with the witness δ_0, its bound on line 5 and the instance on line 7.
- Line 9 ($\exists$E, flag δ_0): δ_0 does not appear in line 9 (the witness was forgotten at line 8), in line 4, in the premise, or in the assumptions still open at line 9, lines 2 and 3.
- Line 10 ($\forall$I, flag ε_0): ε_0 does not appear in the premise, in line 10, or in the assumption still open at line 10, line 2.
- Line 11 ($\forall$I, flag x_0): x_0 does not appear in the premise or in line 11, and no assumption is open.

The answer to the prompt: line 5 is the assumption of the $\exists$E box, so it rests on line 4, which rests on lines 1 and 3. None of the three mentions x_0. The witness δ_0 was obtained from the premise at ε_0 alone, before x_0 played any part. That is uniformity, read off the lines.

The derivation has 11 lines. Written with the unbounded rules it has 17, since each of its six

bounded steps (lines 4, 7, 8, 9, 10 and 11) costs one line more (Section 6.1). The bounded rules keep it at the level at which a mathematician thinks, and the same argument in prose is three sentences, each aligned with the lines it abbreviates:

> Let $x \in A$ and $\varepsilon > 0$ [lines 2, 3, 10, 11]. By uniform continuity there is $\delta > 0$ such that $|u - v| < \delta$ implies $|f(u) - f(v)| < \varepsilon$ for all $u, v \in A$; choose such a δ [4, 5, 6, 9]. Applied with $u = x$, this gives $|f(x) - f(y)| < \varepsilon$ for every $y \in A$ with $|x - y| < \delta$, so this δ works for x and ε [7, 8].

The prose renames the bound variables of UC to u, v, so that "x" can name the fixed point; a derivation needs no such care, because ∀E handles bound variables by substitution.

6.4.4 Reading dependence off the boxes

The derivation shows something the informal proof leaves unsaid: the exact place where δ is chosen, and what surrounds it there.

A constant chosen in a box may depend on the names of the boxes open around that box, and it does depend only on the names in the line it comes from.

Quick check. In the derivation above, which flagged boxes surround the box of δ_0, opened at line 5? What may δ_0 therefore depend on, and what does it in fact depend on?

Solution. The boxes of x_0 (opened at line 2) and of ε_0 (opened at line 3). So δ_0 may depend on x_0 and on ε_0, as C allows. In fact it depends only on ε_0: it comes from line 4, which mentions ε_0 and not x_0.

Why "may"? Inside the boxes of x_0 and ε_0, every term built from x_0 and ε_0 is available, so a premise could have been instantiated at such a term, and a witness obtained from the resulting ∃ line would depend on both names. The side conditions permit that here, because the goal C puts $\exists\delta$ inside both ∀'s: the witness is forgotten by ∃I (line 8) before either box closes. Why "does"? Because everything known about δ_0 is what its ∃ line says, and line 4 mentions ε_0 alone.

Second reading · *Choosing the delta before the point*

So δ_0 could have been obtained before x_0 entered at all: fix ε_0, obtain δ_0, and only then let $x_0 \in A$. That is the order of UC itself. A derivation in that order nests the boxes as ε_0, δ_0, x_0 (the box of δ_0 rests on line 4, so it stays inside the box of ε_0), and it ends in $\forall\varepsilon > 0\ \forall x \in A\ \exists\delta > 0\ \forall y \in A\ (\dots)$, which is C with its first two quantifiers exchanged. C follows by exchanging them back, as in the swap of Example (Flagged boxes at work) (c), here with the bounded rules: boxes for $x_1 \in A$ and $\varepsilon_1 > 0$, ∀E at ε_1 and then at x_1, and two ∀I lines. C's order of choices allows more dependence than this proof uses.

The limit derivation of Section 6.1 showed the same thing on a smaller scale: its witness N_0 came from a line obtained at $\varepsilon_0/2$, inside the box of ε_0 only, and so it depended on ε_0.

6.4.5 The converse breaks at a side condition

The converse, from C to UC, is the interesting direction. The search for it fails, and the failure has an exact address.

Example. The search from C to UC

Run the search method on the sequent with premise C and goal UC, and find where it stops.

Try it first (ten minutes): start the search. Where does a δ come from, and at what moment can you get one?

Idea. The obvious plan copies the derivation above: fix ε_0, obtain δ_0 from the premise, choose it as the witness. It runs into an obstacle at the second step. In C, $\exists\delta$ stands behind $\forall x \in A$, so to reach it the premise must first be instantiated at a point, and the bounded $\forall$E needs a line $t \in A$. No such line exists until we assume one, in a box for some x_0. The δ_0 obtained inside that box comes from a line that mentions x_0, and both side conditions then hold it in: it cannot leave its own box while a line about it still mentions it, and x_0 cannot be generalized while a line about δ_0 still mentions x_0.

Solution. The goal UC begins $\forall\varepsilon > 0$, so open a box for ε_0. The new goal $\exists\delta > 0\ \forall x \in A\ \forall y \in A\ (\dots)$ is postponed. The premise C begins $\forall x \in A$, and no line $t \in A$ is available; the only way to get one is to assume it, so open a bounded box "let $x_0 \in A$". Then the premise yields a δ. Here is how far the search gets, every line legal:

1	$\forall x \in A\ \forall\varepsilon > 0\ \exists\delta > 0\ \forall y \in A\ (\lvert x-y\rvert < \delta \to \lvert f(x)-f(y)\rvert < \varepsilon)$	
2	ε_0 | $\varepsilon_0 > 0$	asm, for $\forall$I
3	x_0 | $x_0 \in A$	asm, for $\forall$I
4	$\forall\varepsilon > 0\ \exists\delta > 0\ \forall y \in A\ (\lvert x_0-y\rvert < \delta \to \lvert f(x_0)-f(y)\rvert < \varepsilon)$	$\forall$E 1, 3
5	$\exists\delta > 0\ \forall y \in A\ (\lvert x_0-y\rvert < \delta \to \lvert f(x_0)-f(y)\rvert < \varepsilon_0)$	$\forall$E 4, 2
6	δ_0 | $\delta_0 > 0$	asm, for $\exists$E
7	$\forall y \in A\ (\lvert x_0-y\rvert < \delta_0 \to \lvert f(x_0)-f(y)\rvert < \varepsilon_0)$	asm, for $\exists$E
8	$\exists\delta > 0\ \forall y \in A\ (\lvert x_0-y\rvert < \delta \to \lvert f(x_0)-f(y)\rvert < \varepsilon_0)$	$\exists$I 6, 7
9	$\exists\delta > 0\ \forall y \in A\ (\lvert x_0-y\rvert < \delta \to \lvert f(x_0)-f(y)\rvert < \varepsilon_0)$	$\exists$E 5, 6–8
10	$\forall x \in A\ \exists\delta > 0\ \forall y \in A\ (\lvert x-y\rvert < \delta \to \lvert f(x)-f(y)\rvert < \varepsilon_0)$	$\forall$I 3–9
11	$\forall\varepsilon > 0\ \forall x \in A\ \exists\delta > 0\ \forall y \in A\ (\lvert x-y\rvert < \delta \to \lvert f(x)-f(y)\rvert < \varepsilon)$	$\forall$I 2–10

The goal inside the box of ε_0 needs a δ that works for **all** points: the line $\forall x \in A\ \forall y \in A\ (\lvert x-y\rvert < \delta_0 \to \lvert f(x)-f(y)\rvert < \varepsilon_0)$, so that $\exists$I at δ_0 could finish. Line 7 says this only for the point x_0. There are two ways to try to generalize it, and each is stopped by a side condition.

- **Generalize inside the box of δ_0.** After line 7, open a new flagged box with the name x_0 and the assumption $x_0 \in A$, repeat line 7 in it, and close it by $\forall$I with the

conclusion $\forall x \in A\ \forall y \in A\ (|x - y| < \delta_0 \to \dots)$. That line breaks the side condition of ∀I: x_0 appears in assumptions still open there, line 3 and line 7. (With a different name x_1 the box is legal, but line 7 says nothing about x_1, so the box cannot reach its last line.)

- **Leave the box of δ_0 first, carrying line 7.** Close the ∃E box with line 7 itself as its conclusion, in order to generalize over x_0 outside it. That line breaks the side condition of ∃E: its conclusion mentions the witness δ_0.

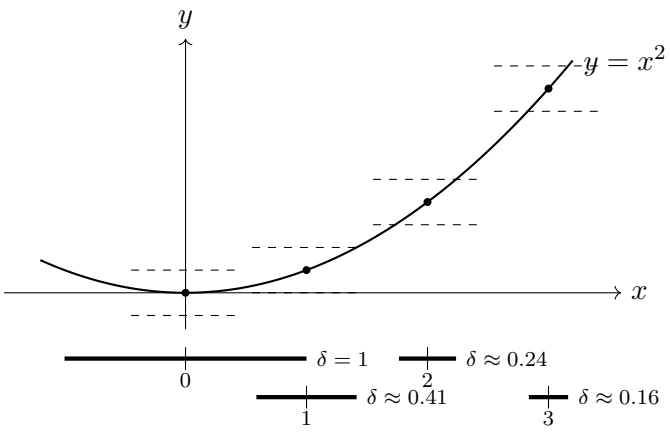
The only legal way out of the box of δ_0 is to forget δ_0, as line 8 does with ∃I before line 9 closes the box. Then ∀I on x_0 is legal, and line 10 is what it gives: $\forall x \in A\ \exists \delta > 0$, with δ still behind x. That is C at ε_0 again, not UC, and closing the box of ε_0 gives line 11, which is C with its first two quantifiers exchanged. The search has gone round in a circle.

These are the same two blocks that stopped the converse of Proposition 5.6.2: a ∀I blocked by an open assumption that mentions its name, and an ∃E box that cannot close on a line mentioning its witness. That is no coincidence: with the bounds and the matrix stripped away, C to UC is the swap $\forall x\, \exists \delta$ to $\exists \delta\, \forall x$.

A failed search proves nothing about all derivations; a counter-reading does, and the next subsection gives one: a function for which C is true and UC false. Section 7.2 explains why that settles the question.

6.4.6 A function that is continuous and not uniformly continuous

Here is the picture first; the example proves it.



With $\varepsilon = 1$, the largest δ that works at $x \geq 0$ is $\sqrt{x^2 + 1} - x$: 1 at 0, about 0.41 at 1, 0.24 at 2, 0.16 at 3. It shrinks as x grows, and no single positive δ works at every x.

Example. The square function on the real line

Let $f(x) = x^2$ on $A = \mathbb{R}$. Show that f satisfies C and does not satisfy UC.

Solution. C holds. Let $x \in \mathbb{R}$ and $\varepsilon > 0$. Let

$$\delta := \min\Big(1,\ \frac{\varepsilon}{2|x|+1}\Big),$$

which is positive, since $\varepsilon > 0$ and $2|x|+1 \geq 1 > 0$. It depends on ε **and on** x. Let $y \in \mathbb{R}$ with $|x-y| < \delta$. By the triangle inequality (Example (The triangle inequality by cases)), used twice, and since $\delta \leq 1$, $|y| \leq |x| + |y-x| < |x|+1$, so $|x+y| \leq |x| + |y| < 2|x|+1$. Hence

$$|x^2 - y^2| = |x-y| \cdot |x+y| \leq |x-y|\,(2|x|+1) < \delta\,(2|x|+1) \leq \varepsilon,$$

where the strict step uses $2|x|+1 > 0$, and the last step uses $\delta \leq \varepsilon/(2|x|+1)$. So this δ works for x and ε.

UC fails. By the negation of uniform continuity in Example (Three negations from analysis) (c), with the bounds $x \in \mathbb{R}$ and $y \in \mathbb{R}$ dropped because they always hold, we must show

$$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x\ \exists y\ (|x-y| < \delta \wedge |x^2 - y^2| \geq \varepsilon).$$

Take $\varepsilon = 1$. Let $\delta > 0$ be arbitrary, and take $x := 1/\delta$ and $y := 1/\delta + \delta/2$, which depend on δ. Then $|x-y| = \delta/2 < \delta$, and, since $y > x > 0$ (so $y-x$ and $y+x$ are both positive),

$$|x^2 - y^2| = (y-x)(y+x) = \frac{\delta}{2}\Big(\frac{2}{\delta} + \frac{\delta}{2}\Big) = 1 + \frac{\delta^2}{4} \geq 1.$$

Since δ was arbitrary, UC fails.

The δ for C grows smaller as $|x|$ grows, and no single δ serves the whole line. The proof of failure says why: however small δ is, far enough out (at $x = 1/\delta$) the graph is so steep that two points $\delta/2$ apart have values more than 1 apart. Read in the real numbers, with f as squaring, the premise C is true and the conclusion UC is false: a counter-reading, and by Section 7.2 no derivation of UC from C exists.

Warning. **"Depending on ε" is a claim, and it can be false.** When a proof says "choose δ (depending on ε)", the parenthesis records which boxes are open where δ is chosen, and which names its formula uses. This is the logical origin of the book's rule that every chosen constant says what it depends on: the dependence is a record of the open boxes, and a reader can check it. A proof of uniform continuity must be able to write "depending on ε" truthfully. A "proof" that x^2 is uniformly continuous on $\mathbb{R}$ that takes $\delta = \min(1, \varepsilon/(2|x|+1))$ "(depending on ε)" has a formula for δ that mentions x: the δ was chosen inside the box of x, so the proof shows only C, and Example (The square function on the real line) shows that UC is false.

The proof that UC fails for x^2 established, with numbers, the positive form of the negation from Section 6.2. That form, taken before its last order fact, refutes UC by logic alone, and a derivation shows how; one TAUT line (Section 6.3) carries its propositional core.

Example. The negation at work, in one derivation

Let E be $\exists \varepsilon > 0\ \forall \delta > 0\ \exists x \in A\ \exists y \in A\ (|x-y| < \delta \wedge \neg(|f(x) - f(y)| < \varepsilon))$, the form of Example (Three negations from analysis) (c) before its last order fact. Derive $E \vdash \neg$UC,

with the bounded rules of Section 6.1.

Solution.

1		$\exists \varepsilon > 0\ \forall \delta > 0\ \exists x \in A\ \exists y \in A\ (\lvert x - y\rvert < \delta \wedge \neg(\lvert f(x) - f(y)\rvert < \varepsilon))$	
2		$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A\ \forall y \in A\ (\lvert x - y\rvert < \delta \to \lvert f(x) - f(y)\rvert < \varepsilon)$	asm, for ¬I
3	ε_1	$\varepsilon_1 > 0$	asm, for ∃E
4		$\forall \delta > 0\ \exists x \in A\ \exists y \in A\ (\lvert x - y\rvert < \delta \wedge \neg(\lvert f(x) - f(y)\rvert < \varepsilon_1))$	asm, for ∃E
5		$\exists \delta > 0\ \forall x \in A\ \forall y \in A\ (\lvert x - y\rvert < \delta \to \lvert f(x) - f(y)\rvert < \varepsilon_1)$	∀E 2, 3
6	δ_1	$\delta_1 > 0$	asm, for ∃E
7		$\forall x \in A\ \forall y \in A\ (\lvert x - y\rvert < \delta_1 \to \lvert f(x) - f(y)\rvert < \varepsilon_1)$	asm, for ∃E
8		$\exists x \in A\ \exists y \in A\ (\lvert x - y\rvert < \delta_1 \wedge \neg(\lvert f(x) - f(y)\rvert < \varepsilon_1))$	∀E 4, 6
9	x_1	$x_1 \in A$	asm, for ∃E
10		$\exists y \in A\ (\lvert x_1 - y\rvert < \delta_1 \wedge \neg(\lvert f(x_1) - f(y)\rvert < \varepsilon_1))$	asm, for ∃E
11	y_1	$y_1 \in A$	asm, for ∃E
12		$\lvert x_1 - y_1\rvert < \delta_1 \wedge \neg(\lvert f(x_1) - f(y_1)\rvert < \varepsilon_1)$	asm, for ∃E
13		$\forall y \in A\ (\lvert x_1 - y\rvert < \delta_1 \to \lvert f(x_1) - f(y)\rvert < \varepsilon_1)$	∀E 7, 9
14		$\lvert x_1 - y_1\rvert < \delta_1 \to \lvert f(x_1) - f(y_1)\rvert < \varepsilon_1$	∀E 13, 11
15		$\bot$	TAUT 12, 14
16		$\bot$	∃E 10, 11–15
17		$\bot$	∃E 8, 9–16
18		$\bot$	∃E 5, 6–17
19		$\bot$	∃E 1, 3–18
20		$\neg\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A\ \forall y \in A\ (\lvert x - y\rvert < \delta \to \lvert f(x) - f(y)\rvert < \varepsilon)$	¬I 2–19

Line 15's propositional skeleton is $p \wedge \neg q,\ p \to q \models \bot$, so the line replaces ∧E twice, →E and ¬E, four lines. The boxes on the witness lines 5, 8 and 10 are opened as soon as those lines exist, the move "name the witness early" (Section 5.6), which Section 6.3 puts first. Every closing line's witness (y_1, x_1, δ_1, ε_1) is absent from $\bot$, from the ∃ line it cites, from the premise, and from the assumptions still open.

6.4.7 Witness functions

UC says that for every $\varepsilon > 0$ some $\delta > 0$ works for all points. A rule that picks such a δ for each ε is a function, $\delta(\varepsilon)$: a **witness function** for UC. For $f(x) = x/2$ on $\mathbb{R}$, $\delta(\varepsilon) = 2\varepsilon$ works, since $\lvert x/2 - y/2\rvert = \lvert x - y\rvert/2 < \varepsilon$ when $\lvert x - y\rvert < 2\varepsilon$.

For C the witness may depend on the point too: $\delta(x, \varepsilon)$. For x^2, Example (The square function

on the real line) gave $\delta(x, \varepsilon) = \min(1, \varepsilon/(2|x| + 1))$.

A derivation shows which arguments a witness function needs. A witness obtained inside the boxes of ε_0 and x_0 may be a function of both, and one obtained inside the box of ε_0 alone is a function of ε_0 alone. "May depend" and "does depend" are exactly the two readings of "Reading dependence off the boxes". So UC holds when some witness function for C ignores its point argument, and the derivation of Example (Uniformly continuous implies continuous) does the reverse: it turns a $\delta(\varepsilon)$ for UC into a $\delta(x, \varepsilon) := \delta(\varepsilon)$ for C, one that ignores x.

In the game of Section 5.1, a witness function is a strategy for the defender: a rule that answers every challenge, and the defender wins by following it.

A formula of this book cannot say "there is a function": its variables range over numbers. A function symbol can stand for one chosen function, and that is how Sections 5.1 and 6.1 wrote sequences.

A tempting "proof" of the converse tries to build a witness function for UC out of one for C, by taking the smallest of its values.

Example. A "proof" that continuity gives uniform continuity

Find the error in this argument, which claims that every function continuous at every point of A is uniformly continuous on A.

> *Let $\varepsilon > 0$. For each $x \in A$, continuity at x gives $\delta_x > 0$ with $|f(x) - f(y)| < \varepsilon$ whenever $y \in A$ and $|x - y| < \delta_x$. Let δ be the smallest of the numbers δ_x. Then $\delta > 0$, and it works at every point, so f is uniformly continuous.*

Solution.

- The conclusion is false for x^2 on $\mathbb{R}$ (Example (The square function on the real line)), so some step fails, and it is the third sentence.
- Infinitely many positive numbers need not have a smallest, and they may come arbitrarily close to 0. With $\varepsilon = 1$, the largest δ that works for x^2 at $x \geq 0$ is $\sqrt{x^2 + 1} - x$ (the figure above), which is $1/(\sqrt{x^2 + 1} + x) < 1/(2x)$ for $x > 0$, and so smaller than any given positive number once x is large.
- In the rules: each δ_x is the witness of an $\exists$E box inside the box of x_0, and "the smallest of them" would be a line about all the witnesses at once. Using it needs a δ that works at every point, the line $\forall x \in A\ \forall y \in A\ (|x - y| < \delta \to \dots)$ whose two routes Example (The search from C to UC) found blocked by side conditions.
- For finite $A = \{a_1, \dots, a_k\}$ the proof is right: k uses of $\exists$E give $\delta_1, \dots, \delta_k$ (the subsection "Two witnesses" of Section 5.6, repeated), and $\min(\delta_1, \dots, \delta_k) > 0$ works at every point.

The search stalls at the same side condition that stopped Example (The search from C to UC); that x^2 is a counter-reading for the converse is the evidence that no other route exists (Section 7.2).

Quick check. For which sets A does the step "let δ be the smallest of the δ_x" make sense, and give a positive δ?

Solution. For finite non-empty A. Finitely many positive numbers have a smallest, and it is positive. For empty A, UC holds outright, since its $\forall x \in A$ is vacuous.

Remark (Looking ahead). Chapter 8 shows how the eleven lines of Example (Uniformly continuous implies continuous) become the three sentences of a written proof, at several levels of detail.

6.4.8 Exercises

6.4.8.1 A. Check your understanding

Exercise. A1

These parts are about the derivation of Example (Uniformly continuous implies continuous).

(a) Why is line 7 legal? Name the rule, the term, and the lines it cites, and say why those lines are available.

(b) At which line is the premise's "for all $x \in A$" used, and at which term?

(c) State the idea of the derivation in two sentences.

(d) Run the derivation for $f(x) = 3x + 1$ on $A = \mathbb{R}$, where UC holds with $\delta = \varepsilon/3$. For a given x_0 and $\varepsilon_0 > 0$, which δ does the derivation hand to C at line 8, and does it depend on x_0? Check that it works.

Solution. *(a)* Line 7 is the bounded ∀E at the term x_0. It cites line 6, $\forall x \in A\ \forall y \in A\ (\dots)$, and line 2, $x_0 \in A$, the bound's instance. Line 6 is the assumption of the ∃E box that contains line 7, and line 2 is the assumption of the box of x_0, which also contains line 7; both boxes are still open at line 7. The term x_0 is a variable not appearing in line 6, so it is free for x.

(b) At line 7, at the term x_0. Line 6 says that δ_0 works for **all** pairs of points of A, and line 7 uses that for pairs with first point x_0.

(c) Uniform continuity at ε supplies one δ that works for all points at once, and it is supplied before any point is named. Continuity at a point x then takes that same δ, specialized to the point x.

(d) Line 4 says that some $\delta > 0$ works at ε_0, and line 8 hands C that witness δ_0, whatever it is; the derivation knows only lines 5 and 6 about it. Here $\varepsilon_0/3$ is such a witness (so is any δ with $0 < \delta \leq \varepsilon_0/3$). Reading the derivation with the witness the exercise names, $\delta_0 = \varepsilon_0/3$, the δ handed to C is $\varepsilon_0/3$. It does not depend on x_0. It works: for every real y with $|x_0 - y| < \varepsilon_0/3$,

$$|f(x_0) - f(y)| = |(3x_0 + 1) - (3y + 1)| = 3|x_0 - y| < 3 \cdot \frac{\varepsilon_0}{3} = \varepsilon_0.$$

Exercise. A2: Drill

Each part takes a line or two. The parts run from single steps, through a missing line, a justification column and a step to judge, to witness functions, a negation and one short derivation. UC and C are the formulas of "Two definitions that differ in one place".

(a) Line 1 is UC and line 2 is $1 > 0$. Write the line that the bounded $\forall$E 1, 2 gives.

(b) The available lines are $\delta_1 > 0$ and $\forall y \in A\ (|x_0 - y| < \delta_1 \to |f(x_0) - f(y)| < \varepsilon_0)$. Write the line that the bounded $\exists$I with witness δ_1 gives.

(c) In Example (The negation at work, in one derivation), line 13 is the bounded $\forall$E from line 7 at x_1. Which second line must be available?

(d) In the six-line derivation of "The skeleton first", give the rule and the cited lines of lines 3 and 4.

(e) In Example (The search from C to UC), give the rule and the cited lines of line 10.

(f) Is this correct? Inside the boxes of ε_0 and x_0, a witness δ_0 obtained from C at x_0 is used in $\exists$I to write $\exists\delta > 0\ \forall x \in A\ \forall y \in A\ (\dots)$.

(g) Give a witness function for UC of $f(x) = 5x$ on $\mathbb{R}$.

(h) For x^2, what does the witness function $\min(1, \varepsilon/(2|x| + 1))$ give at $x = 3$, $\varepsilon = 1$?

(i) For x^2 and $\varepsilon = 1$, what is the largest δ that works at $x = 4$, by the formula under the figure?

(j) Negate C, "continuous at every point of A", into positive form.

(k) Which quantifier of C must move, and past which, to turn C into UC?

(l) Derive $\forall\varepsilon > 0\ \exists\delta > 0\ (\delta \le \varepsilon)$ with the bounded rules, given the arith line $\varepsilon_0 \le \varepsilon_0$.

Solution.

(a) $\exists\delta > 0\ \forall x \in A\ \forall y \in A\ (|x - y| < \delta \to |f(x) - f(y)| < 1)$.

(b) $\exists\delta > 0\ \forall y \in A\ (|x_0 - y| < \delta \to |f(x_0) - f(y)| < \varepsilon_0)$.

(c) Line 9, $x_1 \in A$.

(d) Line 3: $\forall$E 2, at x_0. Line 4: $\exists$I 3, with witness δ_0.

(e) The bounded $\forall$I, 3–9.

(f) No. That $\exists$I needs the line $\forall x \in A\ \forall y \in A\ (|x - y| < \delta_0 \to \dots)$, and that line needs $\forall$I over x_0 while an open assumption about δ_0 mentions x_0: the side condition of $\forall$I.

(g) $\delta(\varepsilon) = \varepsilon/5$, since $|5x - 5y| = 5|x - y| < \varepsilon$ when $|x - y| < \varepsilon/5$.

(h) $\min(1, 1/7) = 1/7$.

(i) $\sqrt{17} - 4 \approx 0.123$.

(j) $\exists x \in A\ \exists\varepsilon > 0\ \forall\delta > 0\ \exists y \in A\ (|x - y| < \delta \wedge |f(x) - f(y)| \ge \varepsilon)$.

(k) $\forall x \in A$ moves to the right, past $\exists\delta > 0$. Passing $\forall\varepsilon > 0$ is harmless (two $\forall$'s exchange); passing $\exists\delta$ is the swap that fails.

(l) The witness function is $\delta(\varepsilon) = \varepsilon$:

1	ε_0 | $\varepsilon_0 > 0$		asm, for $\forall$I
2	| $\varepsilon_0 \le \varepsilon_0$		arith
3	| $\exists\delta > 0\ (\delta \le \varepsilon_0)$		$\exists$I 1, 2
4	$\forall\varepsilon > 0\ \exists\delta > 0\ (\delta \le \varepsilon)$		$\forall$I 1–3

6.4.8.2 B. Practice

Exercise. B1: Lipschitz implies uniformly continuous

A function f is **Lipschitz** on A when

$$\exists L \geq 0\ \forall x \in A\ \forall y \in A\ (|f(x) - f(y)| \leq L \cdot |x - y|).$$

Prove that a function that is Lipschitz on A is uniformly continuous on A, first as a derivation with the bounded rules (lines of arithmetic justified "arith"), then as a short proof in prose. Say what your δ depends on. (The bound $L \geq 0$ is part of the definition: if A has at most one point, then $x = y$ in the matrix, and any L, even $L = -1$, would satisfy it. The bound rules this out, so a proof may use $L \geq 0$, for instance to divide by $L + 1$.)

Solution. The goal UC begins $\forall \varepsilon > 0$: open a box for ε_0. The new goal $\exists \delta > 0 \ldots$ is postponed until a witness is in hand. The premise begins $\exists L \geq 0$: open its $\exists$E box at once, with a witness L_0, aiming for the postponed goal, which does not mention L_0. Work backwards from what the boxes for x_0 and y_0 will need: from $|x_0 - y_0| < \delta$ we want $|f(x_0) - f(y_0)| < \varepsilon_0$, and the premise gives $|f(x_0) - f(y_0)| \leq L_0|x_0 - y_0| < L_0\delta$ when $L_0 > 0$. So $\delta = \varepsilon_0/L_0$ is the natural guess, but $L_0 = 0$ is allowed and then ε_0/L_0 is undefined. Take $\delta = \varepsilon_0/(L_0 + 1)$ instead: $L_0 + 1 \geq 1 > 0$, and $L_0\delta = \varepsilon_0 \cdot L_0/(L_0 + 1) < \varepsilon_0$.

1		$\exists L \geq 0\ \forall x \in A\ \forall y \in A\ (\lvert f(x) - f(y)\rvert \leq L \cdot \lvert x - y\rvert)$	
2	ε_0	$\varepsilon_0 > 0$	asm, for $\forall$I
3	L_0	$L_0 \geq 0$	asm, for $\exists$E
4		$\forall x \in A\ \forall y \in A\ (\lvert f(x) - f(y)\rvert \leq L_0 \cdot \lvert x - y\rvert)$	asm, for $\exists$E
5		$\varepsilon_0/(L_0{+}1) > 0$	arith 2, 3
6	x_0	$x_0 \in A$	asm, for $\forall$I
7	y_0	$y_0 \in A$	asm, for $\forall$I
8		$\lvert x_0 - y_0\rvert < \varepsilon_0/(L_0{+}1)$	asm, for $\to$I
9		$\forall y \in A\ (\lvert f(x_0) - f(y)\rvert \leq L_0 \cdot \lvert x_0 - y\rvert)$	$\forall$E 4, 6
10		$\lvert f(x_0) - f(y_0)\rvert \leq L_0 \cdot \lvert x_0 - y_0\rvert$	$\forall$E 9, 7
11		$\lvert f(x_0) - f(y_0)\rvert < \varepsilon_0$	arith 2, 3, 8, 10
12		$\lvert x_0 - y_0\rvert < \varepsilon_0/(L_0{+}1) \to \lvert f(x_0) - f(y_0)\rvert < \varepsilon_0$	$\to$I 8–11
13		$\forall y \in A\ (\lvert x_0 - y\rvert < \varepsilon_0/(L_0{+}1) \to \lvert f(x_0) - f(y)\rvert < \varepsilon_0)$	$\forall$I 7–12
14		$\forall x \in A\ \forall y \in A\ (\lvert x - y\rvert < \varepsilon_0/(L_0{+}1) \to \lvert f(x) - f(y)\rvert < \varepsilon_0)$	$\forall$I 6–13
15		$\exists \delta > 0\ \forall x \in A\ \forall y \in A\ (\lvert x - y\rvert < \delta \to \lvert f(x) - f(y)\rvert < \varepsilon_0)$	$\exists$I 5, 14
16		$\exists \delta > 0\ \forall x \in A\ \forall y \in A\ (\lvert x - y\rvert < \delta \to \lvert f(x) - f(y)\rvert < \varepsilon_0)$	$\exists$E 1, 3–15
17		$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A\ \forall y \in A\ (\lvert x - y\rvert < \delta \to \lvert f(x) - f(y)\rvert < \varepsilon)$	$\forall$I 2–16

Line 11 is arithmetic from lines 2, 3, 8 and 10: $|f(x_0) - f(y_0)| \leq L_0|x_0 - y_0| \leq L_0 \cdot \varepsilon_0/(L_0 + 1)$,

where the second step uses $L_0 \geq 0$, and $L_0/(L_0+1) < 1$ with $\varepsilon_0 > 0$ gives $L_0\varepsilon_0/(L_0+1) < \varepsilon_0$. Line 15 is the bounded $\exists$I at the term $\varepsilon_0/(L_0+1)$, with its bound on line 5. At line 16, L_0 does not appear in line 16 (the witness term was forgotten at line 15), in line 1, which is the premise, or in the assumption $\varepsilon_0 > 0$ still open there. At lines 13, 14 and 17 the flags y_0, x_0, ε_0 do not appear in the premise, in the line itself, or in the assumptions then open (among lines 2, 3, 4 and 6).

In prose: *Let $L \geq 0$ be a constant with $|f(x) - f(y)| \leq L|x - y|$ for all $x, y \in A$. Let $\varepsilon > 0$, and let $\delta := \varepsilon/(L+1)$, which is positive and depends on ε (and on the constant L, which is fixed before ε), but not on any point. For $x, y \in A$ with $|x-y| < \delta$, $|f(x)-f(y)| \leq L|x-y| \leq L\delta = \varepsilon L/(L+1) < \varepsilon$. So f is uniformly continuous on A.* The box of L_0 is open where δ is chosen, so δ may depend on L_0. That is harmless: L_0 is a constant of the function, not a point, and its box could as well have been opened first, outside the box of ε_0.

Exercise. B2: Put the proof in order

Below are ten sentences. Eight of them, in a suitable order, form a proof that a function f uniformly continuous on A is continuous at every point of A; two are distractors. Write down an order for the eight, and say why each distractor is wrong.

(i) Since x and y lie in A and $|x - y| < \delta$, the choice of δ, applied to $u = x$ and $v = y$, gives $|f(x) - f(y)| < \varepsilon$.

(ii) Let $\delta > 0$ be arbitrary.

(iii) Let $\varepsilon > 0$.

(iv) Since $\varepsilon > 0$ and $x \in A$ were arbitrary, f is continuous at every point of A.

(v) Suppose f is uniformly continuous on A.

(vi) By uniform continuity, applied to ε, choose $\delta > 0$ such that for all $u, v \in A$, $|u - v| < \delta$ implies $|f(u) - f(v)| < \varepsilon$.

(vii) Choose $\delta > 0$ for this x and this ε, and then let $x \in A$ be arbitrary.

(viii) Let $x \in A$.

(ix) So for every $y \in A$, $|x - y| < \delta$ implies $|f(x) - f(y)| < \varepsilon$: this δ works for x and ε.

(x) Let $y \in A$ with $|x - y| < \delta$.

Solution. An admissible order: (v), (viii), (iii), (vi), (x), (i), (ix), (iv).

(v) is the hypothesis of the implication being proved, so it opens the proof (its $\to$I box is the outermost one). After it, other orders are admissible as long as each sentence comes after what it uses. (vi) uses the ε of (iii); (x) uses the δ of (vi) and the x of (viii); then (i), (ix), (iv) follow in that order. So exactly three orders are admissible: (viii), (iii), (vi); or (iii), (viii), (vi); or (iii), (vi), (viii), each preceded by (v) and followed by (x), (i), (ix), (iv). Only the first matches the derivation of Example (Uniformly continuous implies continuous) directly. The second and the third fix ε before x; the third even chooses δ before any point is given, which is possible because uniform continuity supplies δ from ε alone. They correspond to variants of that derivation in which the box of ε_0 (in the third, also that of δ_0) is opened outside the box of x_0. Both variants end in $\forall\varepsilon > 0\ \forall x \in A\ \exists\delta > 0\ \forall y \in A\ (\dots)$, and C follows by exchanging the two bounded universal quantifiers (the swap of Example (Flagged boxes at work) (c), with the bounded rules); sentence (iv), "since ε and x were arbitrary", covers both

steps in prose.

Distractor (ii) introduces δ as an arbitrary object, by $\forall$I. But in "continuous at x" the δ is under $\exists$: the proof must **produce** it, by a choice from the hypothesis, not receive it from a challenger. A proof that took δ arbitrary would have to show that every $\delta > 0$ works, which is false in general.

Distractor (vii) chooses δ and only afterwards lets x be arbitrary. It looks legitimate, because when (vii) is read after (viii), x is already fixed, and continuity at x does allow δ to depend on x. But the sentence reverses the order of the two steps: a δ fitted to one particular x is followed by "let x be arbitrary", which claims that the same δ works for every x. In the rules: if δ really was chosen for this x, its $\exists$E box assumes a line that mentions x, and generalizing over x while that assumption is open breaks the side condition of $\forall$I, the first blocked line of Example (The search from C to UC). If δ does not depend on x, as here, then "for this x" is idle and the sentence should be (vi).

Exercise. B3: The square root

Prove that $f(x) = \sqrt{x}$ is uniformly continuous on $A = [0, \infty)$. You may use the school inequality $|\sqrt{x} - \sqrt{y}| \le \sqrt{|x - y|}$ for all $x, y \ge 0$. State what your δ depends on.

Solution. Let $\varepsilon > 0$, and let $\delta := \varepsilon^2$, which is positive and depends on ε only. Let $x, y \in [0, \infty)$ with $|x - y| < \delta$. Then, by the inequality given and because the square root is strictly increasing on $[0, \infty)$,

$$|\sqrt{x} - \sqrt{y}| \le \sqrt{|x - y|} < \sqrt{\varepsilon^2} = \varepsilon.$$

Since x and y were arbitrary, this δ works for all pairs of points; since ε was arbitrary, f is uniformly continuous on $[0, \infty)$. The δ was chosen before the points: in a derivation of this proof, its $\exists$I line would lie inside the box of ε_0 only.

Exercise. B4: Witness functions

For each function, give a witness function for uniform continuity on the set named, say what it depends on, and prove that it works.

(a) $f(x) = |x|$ on $\mathbb{R}$.
(b) $f(x) = x^2$ on $[0, 3]$.
(c) Say in one sentence why the formula of (b) gives no witness function on all of $\mathbb{R}$.

Solution. *(a)* $\delta(\varepsilon) = \varepsilon$, which depends on ε only. Let $x, y \in \mathbb{R}$ with $|x - y| < \varepsilon$. By the triangle inequality (Example (The triangle inequality by cases)), $|x| \le |x - y| + |y|$ and $|y| \le |y - x| + |x|$, so $|x| - |y| \le |x - y|$ and $|y| - |x| \le |x - y|$; hence $||x| - |y|| \le |x - y| < \varepsilon$.

(b) $\delta(\varepsilon) = \varepsilon/6$. For $x, y \in [0, 3]$, $|x + y| = x + y \le 6$, so if $|x - y| < \varepsilon/6$, then

$$|x^2 - y^2| = |x - y| \cdot |x + y| \le 6|x - y| < \varepsilon.$$

It depends on ε and on the bound 3 of the set, and on no point.

(c) On $\mathbb{R}$, $|x+y|$ has no bound, so no constant can replace the 6, and Example (The square function on the real line) shows that no witness function exists there at all.

Marking scheme. (10 points) - 3: (a), with $||x|-|y|| \le |x-y|$ proved (for instance by the triangle inequality twice). - 4: (b): 1 for the bound $|x+y| \le 6$, 2 for the chain, 1 for the dependence. - 3: (c). Fatal error (0 for the step): a "witness" that mentions x or y.

6.4.8.3 C. Going deeper

Exercise. C1: Composition, half done

Let f be uniformly continuous on A, with $f(x) \in B$ for every $x \in A$, and let g be uniformly continuous on B. Here is the start of a proof that the composite $x \mapsto g(f(x))$ is uniformly continuous on A:

> Fix $\varepsilon > 0$. By the uniform continuity of g, choose $\eta > 0$ such that $|u-v| < \eta$ implies $|g(u)-g(v)| < \varepsilon$ for all $u, v \in B$.

The rest is blank. *Name the move, then complete it.*

Solution (Hint 1). The goal asks for a δ. Which hypothesis can produce one, and at which positive number should it be applied?

Solution (Hint 2). "By the uniform continuity of f, applied to η, choose $\delta > 0$ such that ..."

Solution. The move is fix, obtain, choose: ε was fixed ($\forall$I), η obtained from the hypothesis on g ($\forall$E at ε, then $\exists$E); now obtain δ from the hypothesis on f at the term the goal needs, η ($\forall$E at η, then $\exists$E), and choose that δ as the witness ($\exists$I).

Completion. Since $\eta > 0$, the uniform continuity of f, applied to η, gives $\delta > 0$ such that $|x-y| < \delta$ implies $|f(x)-f(y)| < \eta$ for all $x, y \in A$; choose such a δ. It depends on η, hence on ε, and on no point. Let $x, y \in A$ with $|x-y| < \delta$. Then $|f(x)-f(y)| < \eta$, and $f(x), f(y) \in B$, so the choice of η, applied to $u = f(x)$ and $v = f(y)$, gives $|g(f(x))-g(f(y))| < \varepsilon$. Since x, y were arbitrary, this δ works for all pairs of points of A; since $\varepsilon > 0$ was arbitrary, $x \mapsto g(f(x))$ is uniformly continuous on A.

Marking scheme. (10 points) - 2: names the move (fix, obtain, choose), or the rules $\forall$E at η, $\exists$E, $\exists$I. - 3: applies the uniform continuity of f at η (not at ε), using $\eta > 0$, and chooses δ before any point. - 3: for arbitrary $x, y \in A$ with $|x-y| < \delta$, gets $|f(x)-f(y)| < \eta$ and then $|g(f(x))-g(f(y))| < \varepsilon$, using $f(x), f(y) \in B$. - 2: concludes with the quantifiers in the order ε, δ, points, and says δ depends on ε only. Fatal errors (0 for the step): δ chosen after x or depending on x; the uniform continuity of g applied to points not known to lie in B.

Exercise. C2: Gluing two sets

Let A and B be sets of real numbers and S the set of numbers that lie in A or in B. Suppose f is uniformly continuous on A and uniformly continuous on B.

(a) Show that f need not be uniformly continuous on S: take $A = (0,1)$, $B = (1,2)$, and $f(x) = 0$ for $x < 1$, $f(x) = 1$ for $x \geq 1$.

(b) Prove that f is uniformly continuous on S when $A = [a,b]$ and $B = [b,c]$ with $a < b < c$.

Solution (Hint 1). For (b), the two sets share the point b. Two points on opposite sides of b are each close to b.

Solution (Hint 2). For (b), apply the two hypotheses at $\varepsilon/2$ and take $\delta = \min(\delta_1, \delta_2)$.

Solution. *(a)* On $A = (0,1)$, f is constantly 0, so every $\delta > 0$ works for every $\varepsilon > 0$: $|f(x) - f(y)| = 0 < \varepsilon$. The same holds on $B = (1,2)$, where f is constantly 1. On S we prove the negation of uniform continuity (Example (Three negations from analysis) (c)). Take $\varepsilon = 1$. Let $\delta > 0$, and let $\eta := \min(\delta/4, \frac{1}{2})$, $x := 1 - \eta$, $y := 1 + \eta$. Then $\frac{1}{2} \leq x < 1$, so $x \in A \subseteq S$; $1 < y \leq \frac{3}{2}$, so $y \in B \subseteq S$; $|x - y| = 2\eta \leq \delta/2 < \delta$; and $|f(x) - f(y)| = |0 - 1| = 1 \geq \varepsilon$. So f is not uniformly continuous on S. The gluing fails because the two sets come arbitrarily close at 1 but have no point in common: points close to 1 on opposite sides are each controlled only by their own set, and no common point links them.

(b) Let $\varepsilon > 0$. By the uniform continuity on A at $\varepsilon/2$, choose $\delta_1 > 0$ such that $|f(x) - f(y)| < \varepsilon/2$ for all $x, y \in A$ with $|x - y| < \delta_1$; likewise choose $\delta_2 > 0$ for B. Let $\delta := \min(\delta_1, \delta_2) > 0$, which depends on ε only. Let $x, y \in S$ with $|x - y| < \delta$. There are three cases.

- *Case 1:* $x, y \in A$. Then $|f(x) - f(y)| < \varepsilon/2 < \varepsilon$, since $|x - y| < \delta \leq \delta_1$.
- *Case 2:* $x, y \in B$. The same, with δ_2.
- *Case 3: one point in each set*, say $x \in A$ and $y \in B$ (the other way round is the same with the letters exchanged). Then $x \leq b \leq y$, so $|x - b| = b - x \leq y - x = |x - y| < \delta \leq \delta_1$, and $b \in A$; hence $|f(x) - f(b)| < \varepsilon/2$. Likewise $|b - y| \leq |x - y| < \delta_2$ and $b \in B$, so $|f(b) - f(y)| < \varepsilon/2$. By the triangle inequality (Example (The triangle inequality by cases)),
$$|f(x) - f(y)| \leq |f(x) - f(b)| + |f(b) - f(y)| < \varepsilon.$$

Every point of S lies in A or in B, so the cases cover all pairs. So f is uniformly continuous on S. The common point b is what (a) lacked: there, A and B came arbitrarily close at 1 but had no point in common.

Marking scheme. (10 points) - 3: (a) uniform continuity on A and on B checked, and the negation on S proved: ε first, then for arbitrary δ two points depending on δ, with all three conditions checked. - 2: (b) δ_1 for A at some $\varepsilon_1 > 0$ and δ_2 for B at some $\varepsilon_2 > 0$ with $\varepsilon_1 + \varepsilon_2 \leq \varepsilon$ (the solution takes $\varepsilon_1 = \varepsilon_2 = \varepsilon/2$), and $\delta = \min(\delta_1, \delta_2)$, depending on ε only. - 2: the cases with both points in one set. - 3: the mixed case, through the point b, with $|x - b|, |b - y| \leq |x - y|$ justified and the triangle inequality cited. Fatal errors (0 for the step): a δ depending on x or y; the mixed case omitted; in (a), points chosen before δ.

Exercise. C3: Exit check: a quantifier derivation

(a) Give a full derivation of

$$\exists x\,(P(x) \wedge \forall y\,R(x,y)) \vdash \forall y\,\exists x\,(P(x) \wedge R(x,y)),$$

and check the side condition of every line that closes a flagged box.

(b) For the converse, $\forall y\,\exists x\,(P(x) \wedge R(x,y)) \vdash \exists x\,(P(x) \wedge \forall y\,R(x,y))$, give a domain with two elements and a reading of P and R under which the premise is true and the conclusion false, checking both with the informal reading of Section 5.1.

Solution. (a) The goal is a $\forall$: open a flagged box for y_0. The premise is an $\exists$: open its $\exists$E box at once, with a witness x_0, aiming for the goal's instance. Inside, take the conjunction apart, instantiate $\forall y$ at y_0, and choose x_0 as the witness:

1			$\exists x\,(P(x) \wedge \forall y\,R(x,y))$	
2	y_0	x_0	$P(x_0) \wedge \forall y\,R(x_0,y)$	asm, for $\exists$E
3			$P(x_0)$	$\wedge$E 2
4			$\forall y\,R(x_0,y)$	$\wedge$E 2
5			$R(x_0,y_0)$	$\forall$E 4
6			$P(x_0) \wedge R(x_0,y_0)$	$\wedge$I 3, 5
7			$\exists x\,(P(x) \wedge R(x,y_0))$	$\exists$I 6
8			$\exists x\,(P(x) \wedge R(x,y_0))$	$\exists$E 1, 2–7
9			$\forall y\,\exists x\,(P(x) \wedge R(x,y))$	$\forall$I 2–8

- Line 8 ($\exists$E, flag x_0): x_0 does not appear in line 8, in line 1 (the premise), or in any assumption still open at line 8 (there is none: the box of y_0 has no assumption). Line 7, the last line of the box of x_0, is a line of that box itself and is the formula $\exists x\,(P(x) \wedge R(x,y_0))$ that line 8 repeats (Definition 5.6.1).
- Line 9 ($\forall$I, flag y_0): y_0 does not appear in the premise or in line 9, and no assumption is open. Line 8, the last line of the box of y_0, is the instance at y_0 of $\exists x\,(P(x) \wedge R(x,y))$.

The other nesting, with the box of y_0 inside the box of x_0, works too: $\forall$I at the end of the inner box (its name y_0 is in no open assumption, since the assumption on line 2 does not mention it), then $\exists$E on a conclusion without x_0.

(b) Take a domain with two elements a, b; let P be true of both, and let $R(x,y)$ mean "x is y". The premise is true: for $y = a$ the witness $x = a$ has $P(a)$ and $R(a,a)$, and for $y = b$ the witness $x = b$ works. The conclusion is false: $x = a$ fails $\forall y\,R(a,y)$ at $y = b$, and $x = b$ fails it at $y = a$. The witness in the premise depends on y; the conclusion asks for one x that serves every y. (Section 7.2 makes this informal reading a structure, and shows that it proves the converse is not derivable.)

Marking scheme. (10 points) - 2: a flagged box with a fresh name y_0 for the $\forall$I, closed correctly. - 3: an $\exists$E box with a fresh witness x_0, whose conclusion does not contain x_0, with the side condition checked. - 2: $\wedge$E, $\forall$E at y_0 and $\exists$I with the witness x_0, with the correct terms. - 3: the two-element reading: domain, meaning of P and R, the premise checked for each y and the conclusion refuted for each x. Either nesting of the two boxes earns full marks. Fatal errors (0 for the step): $\forall$I on a name that appears in an open assumption; an

∃E conclusion containing the witness.

6.4.8.4 R. Review

Exercise. R1

Derive $\forall x\,(f(x) = x) \vdash \forall x\,(f(f(x)) = x)$.

Solution. The goal is a ∀: open a flagged box. The premise is instantiated twice, at the flag and at the compound term $f(x_0)$, the term the goal's instance contains.

1		$\forall x\,(f(x) = x)$	
2	x_0	$f(x_0) = x_0$	∀E 1
3		$f(f(x_0)) = f(x_0)$	∀E 1
4		$f(f(x_0)) = x_0$	=E 2, 3
5		$\forall x\,(f(f(x)) = x)$	∀I 2–4

Line 3 instantiates at $f(x_0)$, which is free for x. Line 4 is transitivity, Proposition 5.7.1 (b), written as its one =E step: the equation of line 2 rewrites the last $f(x_0)$ of line 3. At line 5, x_0 is not in the premise or in line 5, and no assumption is open. The method is Section 5.3's "instantiate at the term you need", here a compound term, with the equality of Section 5.7.

Exercise. R2

Show that the list $P \leftrightarrow \neg Q,\ Q \leftrightarrow \neg R,\ R \leftrightarrow \neg P$ is inconsistent.

Solution. A list is inconsistent when $\bot$ is derivable from it (Definition 3.2.1). The derivation below uses no RAA: three ¬I boxes give $\neg R$, $\neg P$ and $\neg Q$, and each $\leftrightarrow$ E is used in whichever of its two forms (Definition 2.4.1) the line needs.

1	$P \leftrightarrow \neg Q$	
2	$Q \leftrightarrow \neg R$	
3	$R \leftrightarrow \neg P$	
4	$\quad P$	asm, for ¬I
5	$\quad \neg Q$	↔E 1, 4
6	$\quad\quad R$	asm, for ¬I
7	$\quad\quad \neg P$	↔E 3, 6
8	$\quad\quad \bot$	¬E 4, 7
9	$\quad \neg R$	¬I 6–8
10	$\quad Q$	↔E 2, 9
11	$\quad \bot$	¬E 10, 5
12	$\neg P$	¬I 4–11
13	R	↔E 3, 12
14	$\quad Q$	asm, for ¬I
15	$\quad \neg R$	↔E 2, 14
16	$\quad \bot$	¬E 13, 15
17	$\neg Q$	¬I 14–16
18	P	↔E 1, 17
19	$\bot$	¬E 18, 12

A TAUT line (Section 4.8) would do it in one step, since the three formulas have no row that makes them all true: an odd cycle of "is the negation of" is impossible.

Exercise. R3

You have derived $P \wedge (Q \vee R) \vdash (P \wedge Q) \vee (P \wedge R)$. Derive

$$\exists x\,(P(x) \wedge (Q(x) \vee R(x))) \vdash \exists x\,(P(x) \wedge Q(x)) \vee \exists x\,(P(x) \wedge R(x)),$$

and say which line is the propositional law.

Solution.

1		$\exists x\,(P(x) \wedge (Q(x) \vee R(x)))$	
2	x_0	$P(x_0) \wedge (Q(x_0) \vee R(x_0))$	asm, for $\exists$E
3		$(P(x_0) \wedge Q(x_0)) \vee (P(x_0) \wedge R(x_0))$	TAUT 2
4		$P(x_0) \wedge Q(x_0)$	asm, for $\vee$E
5		$\exists x\,(P(x) \wedge Q(x))$	$\exists$I 4
6		$\exists x\,(P(x) \wedge Q(x)) \vee \exists x\,(P(x) \wedge R(x))$	$\vee$I 5
7		$P(x_0) \wedge R(x_0)$	asm, for $\vee$E
8		$\exists x\,(P(x) \wedge R(x))$	$\exists$I 7
9		$\exists x\,(P(x) \wedge Q(x)) \vee \exists x\,(P(x) \wedge R(x))$	$\vee$I 8
10		$\exists x\,(P(x) \wedge Q(x)) \vee \exists x\,(P(x) \wedge R(x))$	$\vee$E 3, 4–6, 7–9
11		$\exists x\,(P(x) \wedge Q(x)) \vee \exists x\,(P(x) \wedge R(x))$	$\exists$E 1, 2–10

Line 3, justified by TAUT on the skeleton $p \wedge (q \vee r) \models (p \wedge q) \vee (p \wedge r)$, is the propositional law Proposition 1.6.1 of Section 1.6; citing it as a derived rule (Definition 2.3.1) gives the same line. At line 11, x_0 is not in line 11 or in line 1, the premise, and no assumption is open. The quantifier steps frame the law: this is the propositional core of Section 6.3.

Chapter 7

Models and Theories

The axioms of a partial order say that $\leq$ is reflexive, antisymmetric and transitive. Do they prove that any two elements are comparable, $\forall x\,\forall y\,(x \leq y \vee y \leq x)$? A search for a derivation stalls, and a stalled search proves nothing. Two dots settle it: a domain $\{a, b\}$ in which each element is $\leq$ itself and nothing else. All three axioms are true there and comparability is false, so, by a statement of Section 7.2, no derivation exists. *What does "true there" mean, and why does one picture rule out every derivation?* This chapter answers both, and then uses such pictures on whole theories: in Section 7.4, "is every group commutative?" is answered by the six symmetries of a triangle.

How to read this chapter. First reading: Sections 7.1, 7.2 and 7.4, and Section 7.5 up to "Isomorphic structures". Second pass: the rest of Section 7.5. Optional: Sections 7.3 and 7.6. The statement that a counter-model rules out every derivation is on the trunk, in Section 7.2; its proof is optional, in Section 7.3. Nothing on the first-reading list uses anything off it.

Before you start: five tempting claims. Each is a claim this chapter corrects; decide each now, and check again at the section named.

Quick check.

1. "$\forall x\,\exists y\,(x < y)$ is true."
2. "A search for a derivation of $\Gamma \vdash \varphi$ stalled, so $\Gamma \nvdash \varphi$."
3. "If σ is true in $(\mathbb{N}, <)$, then σ is derivable."
4. "Every group is commutative: the integers under addition are."
5. "$(\mathbb{N}, <)$ and the even numbers with $<$ can be told apart by some sentence, since one is part of the other."

↻ **Second reading** · *Answers*

1. Not until a structure is fixed: true in $(\mathbb{N}, <)$, false in $(\{1, 2, 3\}, <)$ (Section 7.1).
2. No. A stalled search proves nothing; a counter-model does (Section 7.2).
3. No. Derivable sentences are true in **every** structure; $\forall x\,\exists y\,(x < y)$ is true in $(\mathbb{N}, <)$ and false in a three-element order (Section 7.2).
4. No. The six symmetries of a triangle form a group in which two motions done in the two orders differ (Section 7.4).
5. No. Doubling matches them element for element and order for order, and matched struc-

tures satisfy the same sentences (Section 7.5, second pass).

What you need. Chapters 1–6, above all the semantics of Chapter 4 (rows, $\models$, soundness) and the quantifier rules of Chapter 5. Section 7.5's second half uses functions and bijections in their everyday sense.

Time. About 20 hours on the core track (Sections 7.1, 7.2, 7.4 and the first half of 7.5: all of A, one B and one C per section, the exit checks); about 45 hours for everything.

Minimum before *Sets, Relations, and Cardinality*. Chapters 1–6 and 8, core; Sections 7.1–7.2 before its Chapter 3; Sections 7.4–7.5 recommended.

7.0.1 Roadmap

- **Structures and Satisfaction.** A domain and a meaning for every symbol, an assignment for the free variables, clauses that decide every formula; finite structures drawn; proofs by meaning.
- **Counter-models.** Three pictures, a procedure for finding one, small domains by expansion, each side condition of Chapter 5 broken, and the statement that a counter-model rules out every derivation.
- **Soundness for Predicate Logic** (optional). The proof of that statement, by Chapter 4's induction.
- **Theories and Their Models.** Axioms as sentences, consistency by a model, independence by two models.
- **Definability and Isomorphism.** A formula defines a set; $\exists$ is a projection. Second pass: isomorphic structures satisfy the same sentences; automorphisms move what is not definable.
- **What First-Order Logic Can and Cannot Say** (optional). Completeness, compactness, non-standard numbers, undecidability, all stated without proof.

7.0.2 Named moves

- **Refute with a model.** One structure with the premises true and the conclusion false. (Section 7.2)
- **Prove by meaning, clause by clause.** Unwind the premises and the conclusion through (S1)–(S7). (Section 7.1)
- **Search for a counter-model.** Start from the counterexample set, meet the existential demands first, add only forced elements. (Section 7.2)
- **Expand over a small domain.** On two or three elements a quantifier is a long $\wedge$ or $\vee$. (Section 7.2)
- **Consistent by a model; independent by two models.** (Section 7.4)
- **Define by a formula.** (Section 7.5)
- **An automorphism moves what cannot be defined.** (Section 7.5, second pass)

Where this goes. *Sets, Relations, and Cardinality*, Chapter 1 develops the sets, pairs and functions that Sections 7.1 and 7.5 use in their everyday sense, and its Chapter 3 builds a counter-model by Section 7.2's procedure; its Chapter 2 studies the relations of Section 7.4.

7.0.3 Summary

A structure fixes a domain and a meaning for every symbol, an assignment fixes the free variables, and seven clauses decide every formula (Definition 7.1.3); a sentence is true or false in a structure, never "true" alone. A counter-model makes the counterexample set true; by soundness for predicate logic (stated in Section 7.2, proved in the optional Section 7.3) it rules out every derivation: the stalled searches of Chapters 5 and 6 become theorems of non-derivability, and each side condition of the quantifier rules is shown needed. A theory is a set of axioms (Definition 7.4.1); a model shows it consistent (Proposition 7.4.2), and two models show a sentence independent of it (Proposition 7.4.3), as the antichain does for comparability and the triangle for commutativity. Formulas define sets (Definition 7.5.1), and, on the second pass, isomorphic structures satisfy the same sentences (Theorem 7.5.1).

Proved: the semantics, and every non-derivability result the chapter gets from a counter-model, given the soundness statement. Stated without proof: the undecidability of derivability (Church and Turing; Section 7.2, again in 7.6) and the 2^k bound on counter-models of monadic claims (Section 7.2); in the optional Section 7.6, completeness (every consistent theory has a model), compactness, the Löwenheim–Skolem theorem and the decidability of monadic sentences; and, in the solution to Exercise R2 of Section 7.5, that $P \wedge Q$ cannot be derived from its premise without RAA (nothing depends on it). Beyond this volume: why no consistent theory whose axioms a procedure can list, and which derives the basic facts of arithmetic, decides every sentence about $\mathbb{N}$ (Gödel's incompleteness theorem, named in Section 7.6). Chapter 8 turns derivations into the prose of mathematics.

7.0.4 Terms

- **assignment**: a value for each variable (Definition 7.1.2).
- **automorphism**: an isomorphism onto itself (Definition 7.5.2).
- **consistent, inconsistent theory**: no finite list of axioms derives $\bot$ (Definition 7.4.3).
- **counter-model**: premises true, conclusion false (Definition 7.1.4).
- **definable set**: the elements where a formula holds (Definition 7.5.1).
- **domain**: what a structure's quantifiers range over (Definition 7.1.1).
- **independent sentence**: neither derivable nor refutable (Proposition 7.4.3).
- **isomorphism of structures**: a bijection respecting every symbol (Definition 7.5.2).
- **model of a theory**: a structure with every axiom true (Definition 7.4.2).
- **satisfaction,** $\mathfrak{M}, \ell \models \varphi$: the clauses (S1)–(S7) (Definition 7.1.3).
- **semantic entailment,** $\Gamma \models \varphi$: no counter-model (Definition 7.1.4).
- **structure**: a domain and a meaning for every symbol (Definition 7.1.1).
- **theory,** $T \vdash \sigma$: a set of axioms; derivable from finitely many (Definition 7.4.1).
- **true in a structure,** $\mathfrak{M} \models \sigma$: satisfied under every assignment (Section 7.1).

7.1 Structures and Satisfaction

Sections 5.3 to 6.4 derived a good deal, and some searches stalled. The converse of Proposition 5.6.2 could not be finished: every attempt ran into a side condition. The converse of Proposition 5.6.1 stalled in the same way, and so did the search for $\exists x\, P(x) \vdash P(c)$ in Exercise (C2: A witness is not a name, used later). But a search that fails proves nothing, since a cleverer derivation might exist. Chapter 4 met the same problem for the connectives and solved it with meaning: a row of a truth table that makes the premises true and the conclusion false rules out every derivation at once (Corollary 4.7.3). This section builds the meaning of a quantified formula: a **model**, a domain of objects with a meaning for every symbol. A model

that makes the premises true and the conclusion false is a **counter-model**, the exact form of the counter-readings used since Section 5.3, and Section 7.2 uses it to settle the stalled searches.

Remark (Looking back). Section 4.4 gave each formula of Chapter 4 a truth value by recursion on the formula, one clause per connective, under a row. This section does the same for first-order formulas, with a structure and an assignment in place of a row. Clauses (S3)–(S5) below are Chapter 4's clauses for $\bot$ and the connectives, read as conditions for satisfaction instead of truth under a row; (S6) and (S7), for the quantifiers, are the new ones.

7.1.1 Why a row is not enough

Chapter 4's semantics gave each atom a truth value and computed the rest. Try it on the stalled sequent of Proposition 5.6.2's converse,

$$\forall y\, \exists x\, R(x, y) \vdash \exists x\, \forall y\, R(x, y).$$

Neither side has a main connective among $\neg, \wedge, \vee, \to, \leftrightarrow$: each is a quantified formula, and a truth table would have to treat each as one atom. The sequent would become $p \vdash q$, refuted by the row $p = \mathrm{T}$, $q = \mathrm{F}$. That "refutation" is worthless. The same reading refutes $\forall x\, P(x) \vdash \exists x\, P(x)$, which is derivable (Proposition 5.5.1). Treating a quantified formula as an atom throws away exactly the structure that the quantifier rules use.

So what does make $\forall y\, \exists x\, R(x, y)$ true or false? Section 5.1 read such sentences informally, **over a domain**: "for every y in the domain there is an x in the domain with $R(x, y)$". To decide it we need three things.

- **A domain**: the objects the variables range over. Over $\mathbb{Z}$ the sentence may be true and over a three-element set false.
- **A meaning for each symbol.** R is a predicate symbol, a name with two slots and no meaning of its own (Section 5.2). Read as $>$ on $\mathbb{Z}$, the sentence says "every integer has an integer above it", which is true; read as $<$ on $\mathbb{N}$, it says "every natural number has a natural number below it", which is false at $y = 0$.
- **Values for the free variables**, if there are any. The formula $\exists y\, (y < x)$ is neither true nor false on $\mathbb{N}$ until we know which number x is.

The first two together are a **structure**; the third is an **assignment**. A structure with an assignment decides every formula, by clauses that read the quantifiers as Section 5.1 read them. That is all this section does; Section 7.2 uses the result.

7.1.2 Structures

Situate the notion first. A formula is built from **symbols**: constants such as c and 0, function symbols such as f and $+$, predicate symbols such as P, R and $<$, each function and predicate symbol with its number of argument places, its **arity** (Section 5.2, Definition 5.2.1 and Definition 5.2.2). The symbol $=$ is special: it is part of the logic, with rules of its own (Definition 5.7.1), and it will always mean the same thing. A structure gives every other symbol a meaning.

A structure is a non-empty set of objects together with a meaning for each symbol: an object for each constant, an operation for each function symbol, and a relation for each predicate symbol.

Definition 7.1.1. Structure

Fix a collection of constants, function symbols and predicate symbols, each function and predicate symbol with its arity $n \geq 1$. A **structure** $\mathfrak{M}$ (also called a **model**) for these symbols consists of

(M1) a **non-empty** set D, the **domain** of $\mathfrak{M}$;

(M2) for each constant c, an element $c^{\mathfrak{M}} \in D$;

(M3) for each function symbol f of arity n, a function $f^{\mathfrak{M}} \colon D^n \to D$, defined at **every** n-tuple of elements of D;

(M4) for each predicate symbol R of arity n, a set $R^{\mathfrak{M}}$ of n-tuples of elements of D.

The symbol $=$ is not interpreted by $\mathfrak{M}$: in every structure it means **identity**, "is the same element of D as".

In words. (M1) says what the variables range over, and insists that there is at least one object. (M2) says which object each name names; two constants may name the same object. (M3) makes each function symbol an operation that can be applied to **any** objects of the domain and always returns an object of the domain. (M4) makes each predicate symbol a relation: $R^{\mathfrak{M}}$ is the set of tuples for which R "holds". A one-place predicate symbol P becomes a subset $P^{\mathfrak{M}} \subseteq D$, a two-place one a set of ordered pairs, and so on. Here D^n is the set of n-tuples $(d_1, \ldots, d_n)$ of elements of D; this section uses only their everyday meaning.

We usually describe a structure in one line, naming the domain and the meaning of each symbol, and write "$\mathfrak{M} = (\mathbb{Z}, <)$" for the structure with domain $\mathbb{Z}$ in which the symbol $<$ means the usual order. When a symbol has its usual meaning on a familiar set, we say so and do not write $<^{\mathfrak{M}}$. When the domain is followed by a list of symbols, a semicolon separates the domain from the list, as in $(\mathbb{R}; 1, +, \cdot)$: each listed symbol has its usual meaning on that set, and only the listed symbols belong to the structure. With a single symbol we may also write a comma, as in $(\mathbb{Z}, <)$.

Examples, simplest first.

- *One element.* $D = \{a\}$, with $P^{\mathfrak{M}} = \{a\}$. This is the degenerate structure. Every variable must take the value a, so "for every element" and "there is an element" both mean "a". It will reappear as a test case: a sentence true here and false in some other structure cannot be derivable (by the soundness statement in Section 7.2), and some of the chapter's non-derivable sequents need at least two elements to refute.
- *Two elements, one relation.* $D = \{a, b\}$, with $R^{\mathfrak{M}} = \{(a, a), (b, b)\}$: the relation "is the same as", written out as a set of pairs. A finite structure can be given completely by listing.
- *The integers with their order.* $D = \mathbb{Z}$, with $R^{\mathfrak{M}} = \{(x, y) \in \mathbb{Z}^2 : x > y\}$. Here the relation is described by a condition rather than a list.
- *Arithmetic.* $D = \mathbb{R}$, with the constants 0 and 1 naming zero and one, the function symbols $+$, $\cdot$ and two-place $-$ meaning addition, multiplication and subtraction, the one-place symbols $-$, $(\cdot)^2$ and $|\cdot|$ meaning negation, squaring and absolute value, and $<$, $\leq$, $>$, $\geq$ the usual order. Extended by the other familiar symbols with their usual meanings (numerals, halving, the predicates $\in A$, division with its value at a zero denominator fixed arbitrarily as in the non-example below, and a function symbol for each sequence or function an example uses), this is the structure in which the analysis examples of Chapters 5 and 6 have been read all along.

Second reading · *A non-example, and the repair used for 1/n*

Non-example by minimal change. Take the arithmetic structure and add a one-place function symbol g, meant to be read as $x \mapsto 1/x$. This is **not** a structure: (M3) asks for a function defined at every element of D, and $1/0$ is not defined. Everything else in the description is fine; exactly clause (M3) fails, at the single point 0. There are two repairs. Shrink the domain to $(0, \infty)$, where $1/x$ is always defined; or keep $\mathbb{R}$ and give g any value at 0, say $g^{\mathfrak{M}}(0) := 0$. The second repair is the one this book uses for terms such as $1/n$: the value at 0 is fixed arbitrarily and is never looked at, because such terms are used only under a bound $n \in \mathbb{Z}_{>0}$ (as in Exercise C3 of Section 7.2), exactly as a sequence a_n was read in Section 5.1 as a function symbol whose values off $\mathbb{N}$ are never used.

The domain is non-empty because of Proposition 5.5.1; the subsection "Why domains are non-empty" explains why.

Second reading · *Why total functions, and why identity*

- *Total functions.* Clause (M3) is what makes every term name an object: the term $g(0)$ must have a value, or the formula $g(0) = g(0)$, which =I derives from nothing, would have no truth value.
- *Identity for =.* The rule =E replaces equals by equals **everywhere**. If a structure could read $=$ as some other relation, =E would fail. On $\mathbb{Z}$, read $x = y$ as "x and y have the same parity" and P as "is positive": then $-2 = 2$ and $P(2)$ would both hold, while $P(-2)$, which =E derives from them, fails. With $=$ fixed as identity, no such reading exists.

Warning. A structure must interpret **every** symbol that occurs in the formulas at hand. "Take $D = \mathbb{Z}$" is not yet a structure for $\forall y\, \exists x\, R(x, y)$: until R is given a meaning, the sentence has no truth value, and the same domain makes it true with R read as $>$ and false with R read as "$x \cdot x = y$" (at $y = 2$). Always write the domain **and** the meaning of each symbol.

7.1.3 Drawing a finite structure

A finite structure whose symbols are one- and two-place predicate symbols can be drawn, as Section 5.1 drew its worlds. Each element is a dot. Beside a dot stand its name and the one-place predicates it satisfies. For a two-place relation R, an arrow runs from d to e for each pair (d, e) in $R^{\mathfrak{M}}$, and a loop at d records the pair (d, d). Deciding a sentence in a drawn structure is then reading the picture, dot by dot.

When the relation is a partial order $\leq$, arrows crowd the picture, and a lighter drawing is used. Section 7.4 defines partial orders; here it is enough that $\leq$ is reflexive and transitive and never holds both ways between two different elements. The **Hasse diagram** of such a structure puts d below e when $d \leq e$ and $d \neq e$, and draws a line only for the pairs with nothing strictly between

them. The loops, and the pairs that follow by chaining lines upwards, are understood. In the right picture below, $1 \leq 6$ holds although no line joins 1 and 6: the chain 1, 2, 6 runs upwards.

A picture is a structure only when it says what every symbol means: the caption names the domain and each symbol. Figure F draws two structures. Left: $D = \{a, b, c\}$, $P^{\mathfrak{M}} = \{a\}$, $R^{\mathfrak{M}} = \{(a, b), (b, b), (c, a)\}$. Right: $D = \{1, 2, 3, 6\}$, with $x \leq y$ read as "x divides y".

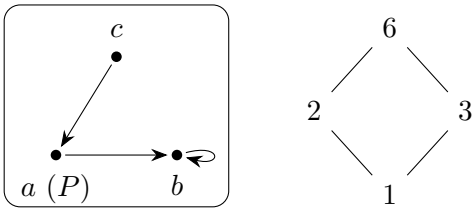


Figure F. An arrow diagram (left) and a Hasse diagram (right).

Quick check. In the left structure of Figure F, decide $\forall x\, \exists y\, R(x, y)$ and $\exists y\, \forall x\, R(x, y)$.

Solution. The first is true: every dot has an arrow out of it ($a \to b$, $b \to b$, $c \to a$). The second is false: no dot receives an arrow from all three. The dot b receives none from c, the dot a none from a or b, and the dot c none at all.

7.1.4 Assignments and the value of a term

A structure names the objects that **constants** stand for. Variables are different: their whole purpose is to range over the domain, so a structure gives them no fixed value. A separate object records what they stand for at a given moment.

Definition 7.1.2. Assignment

Let $\mathfrak{M}$ be a structure with domain D. An **assignment** in $\mathfrak{M}$ is a function ℓ that gives **every** variable x a value $\ell(x) \in D$. For an assignment ℓ, a variable x and an element $d \in D$, the assignment $\ell[x \mapsto d]$ is the one that gives x the value d and every other variable the value ℓ gives it:

$$\ell[x \mapsto d](y) := \begin{cases} d & \text{if } y \text{ is the variable } x, \\ \ell(y) & \text{if } y \text{ is any other variable.} \end{cases}$$

In words: an assignment is a choice of value for every variable at once, and $\ell[x \mapsto d]$ changes that choice at the one variable x, leaving the others alone. The assignment $\ell[x \mapsto d]$ is what the quantifiers will use: "for every x" will mean "for every $d \in D$, with x set to d". Changes can be stacked: $\ell[x \mapsto d][y \mapsto e]$ first sets x to d and then y to e. When x and y are different variables the order does not matter, since each change touches only its own variable; when they are the same variable, the later change wins.

There are always assignments, because D is non-empty: pick any $d_0 \in D$ and let every variable have the value d_0. Like a valuation in Chapter 4 (Definition 4.4.1), an assignment fixes infinitely many values, and a formula will only ever look at finitely many of them.

With a structure and an assignment, every term names an element of the domain. Terms are

built by the clauses of Definition 5.2.1, so the value is defined by recursion on terms, in the pattern of Theorem 4.3.4.

The value of a term is computed from its variables and constants upwards, one function symbol at a time.

For a structure $\mathfrak{M}$ and an assignment ℓ, the **value** $t^{\mathfrak{M},\ell} \in D$ of a term t is:

- $x^{\mathfrak{M},\ell} := \ell(x)$ for a variable x;
- $c^{\mathfrak{M},\ell} := c^{\mathfrak{M}}$ for a constant c;
- $f(t_1,\dots,t_n)^{\mathfrak{M},\ell} := f^{\mathfrak{M}}(t_1^{\mathfrak{M},\ell},\dots,t_n^{\mathfrak{M},\ell})$ for a function symbol f of arity n.

The third clause is where (M3) is needed: $f^{\mathfrak{M}}$ must accept every tuple the terms $t_1,\dots,t_n$ might produce.

For example, in the arithmetic structure on $\mathbb{R}$, let $\ell(x) = 2$ and $\ell(y) = -3$. The term $x \cdot x + y$, which is $+(\cdot(x,x),y)$ written in the usual way, has value $2 \cdot 2 + (-3) = 1$. Under $\ell[y \mapsto 1]$ it has value $2 \cdot 2 + 1 = 5$, and under $\ell[x \mapsto 0]$ it has value -3. The term $1 + 1$ has value 2 under every assignment, since it contains no variable.

7.1.5 Satisfaction

Now the central definition. We want to say when a structure and an assignment make a formula true. Formulas are built by the clauses of Definition 5.2.2, so we define it by recursion, one clause per way of building a formula, and each clause reads its connective or quantifier as Chapter 4 and Section 5.1 read it.

A formula is satisfied when the objects its terms name stand in the relations its predicates name, combined by the connectives as in a truth table, with "for every x" and "there is an x" ranging over the domain.

Definition 7.1.3. Satisfaction

Let $\mathfrak{M}$ be a structure with domain D and ℓ an assignment in $\mathfrak{M}$. The relation $\mathfrak{M},\ell \models \varphi$, read "$\mathfrak{M}$ **satisfies** φ under ℓ", is defined by recursion on the formula φ:

(S1) $\mathfrak{M},\ell \models R(t_1,\dots,t_n)$ if $(t_1^{\mathfrak{M},\ell},\dots,t_n^{\mathfrak{M},\ell}) \in R^{\mathfrak{M}}$;

(S2) $\mathfrak{M},\ell \models t_1 = t_2$ if $t_1^{\mathfrak{M},\ell}$ and $t_2^{\mathfrak{M},\ell}$ are the **same** element of D;

(S3) $\mathfrak{M},\ell \models \bot$ **never** holds;

(S4) $\mathfrak{M},\ell \models \neg\varphi$ if $\mathfrak{M},\ell \models \varphi$ does not hold;

(S5) $\mathfrak{M},\ell \models \varphi \wedge \psi$ if both $\mathfrak{M},\ell \models \varphi$ **and** $\mathfrak{M},\ell \models \psi$; $\mathfrak{M},\ell \models \varphi \vee \psi$ if at least one of them holds; $\mathfrak{M},\ell \models \varphi \to \psi$ unless $\mathfrak{M},\ell \models \varphi$ holds and $\mathfrak{M},\ell \models \psi$ does not; $\mathfrak{M},\ell \models \varphi \leftrightarrow \psi$ if both hold or neither does;

(S6) $\mathfrak{M},\ell \models \forall x\,\varphi$ if $\mathfrak{M},\ell[x \mapsto d] \models \varphi$ **for every** $d \in D$;

(S7) $\mathfrak{M},\ell \models \exists x\,\varphi$ if $\mathfrak{M},\ell[x \mapsto d] \models \varphi$ for **at least one** $d \in D$.

When $\mathfrak{M},\ell \models \varphi$ does not hold we write $\mathfrak{M},\ell \not\models \varphi$. For a list Γ of formulas, $\mathfrak{M},\ell \models \Gamma$ means that $\mathfrak{M},\ell \models \gamma$ for **every** formula γ of Γ.

Read it clause by clause. (S1): an atomic formula is satisfied exactly when the values of its terms form a tuple in the relation that interprets its predicate symbol. (S2): an equation is satisfied when both sides name the same object; this is where "$=$ means identity" enters. (S3)–(S5) are

the $\bot$ half of (V1) and clauses (V2)–(V6) of Chapter 4's Definition 4.4.2, with "is true under v" replaced by "is satisfied by $\mathfrak{M}$ under ℓ": $\bot$ is never satisfied, $\neg$ flips, $\wedge$ needs both, $\vee$ needs one, $\to$ fails only with a satisfied hypothesis and an unsatisfied conclusion, $\leftrightarrow$ needs agreement. (S6) and (S7) are the new clauses. To decide $\forall x\,\varphi$, we run through the elements d of the domain, set x to d and leave every other variable alone, and ask whether φ is satisfied each time. For $\exists x\,\varphi$ one success is enough. In symbols, for (S6) and (S7):

$$\mathfrak{M}, \ell \models \forall x\,\varphi \quad \text{iff} \quad \text{for every } d \in D,\ \ \mathfrak{M}, \ell[x \mapsto d] \models \varphi,$$
$$\mathfrak{M}, \ell \models \exists x\,\varphi \quad \text{iff} \quad \text{there is } d \in D \text{ with } \mathfrak{M}, \ell[x \mapsto d] \models \varphi.$$

A trace, step by step. Decide $\forall x\,\exists y\,(x < y)$ in two structures, both with $<$ the usual order: $\mathfrak{M}_1$ with domain $\mathbb{N}$, and $\mathfrak{M}_2$ with domain $\{1, 2, 3\}$. Let ℓ be any assignment.

In $\mathfrak{M}_1$: by (S6) we must check, for every $d \in \mathbb{N}$, that $\mathfrak{M}_1, \ell[x \mapsto d] \models \exists y\,(x < y)$. Fix d. By (S7) it is enough to find one $e \in \mathbb{N}$ with $\mathfrak{M}_1, \ell[x \mapsto d][y \mapsto e] \models x < y$. Take $e = d + 1$. Under $\ell[x \mapsto d][y \mapsto d+1]$ the terms x and y have the values d and $d+1$, and $d < d+1$, so (S1) holds. Hence the sentence is satisfied. Notice that the witness $e = d + 1$ depends on d: (S7) is checked **inside** (S6), once for each d.

In $\mathfrak{M}_2$: (S6) asks the same for $d = 1, 2, 3$. At $d = 1$ and $d = 2$ the witnesses $e = 2$ and $e = 3$ work. At $d = 3$ no $e \in \{1, 2, 3\}$ has $3 < e$, so (S7) fails for $\ell[x \mapsto 3]$, and (S6) fails. So $\mathfrak{M}_2 \not\models \forall x\,\exists y\,(x < y)$, and the failure has a name: the element 3 has nothing above it.

The assignment matters for free variables. In $\mathfrak{M}_1 = (\mathbb{N}, <)$, with $\leq$ also read as usual, take the formula $\forall y\,(x \leq y)$, in which x is free. Under an assignment with $\ell(x) = 0$ it is satisfied, since every $e \in \mathbb{N}$ has $0 \leq e$ (recall that $\mathbb{N} = \{0, 1, 2, \dots\}$). Under an assignment with $\ell(x) = 1$ it is not: $e = 0$ gives $1 \leq 0$, which fails. The same formula in the same structure is satisfied under one assignment and not under another; it is a property of x, "x is least", not a statement.

↻ **Second reading** · *Satisfaction is well defined; degenerate cases*

Well defined. The clauses define, for each formula, which pairs $(\mathfrak{M}, \ell)$ satisfy it, by recursion on the formula: (S1)–(S3) settle the atomic formulas and $\bot$, and each of (S4)–(S7) settles a compound formula from its immediate parts. One point needs a word. Clauses (S6) and (S7) consult the immediate part φ under **other** assignments $\ell[x \mapsto d]$, not under ℓ. So the recursion must define satisfaction of each formula under **all** assignments at once, and then (S6) reads off the answer for φ at every $\ell[x \mapsto d]$. The recursion is still on the formula alone, and Theorem 4.3.4, applied to formulas built by the clauses of Definition 5.2.2, gives exactly one such relation.

Degenerate cases. In the one-element structure $D = \{a\}$, every assignment gives every variable the value a, so $\ell[x \mapsto a] = \ell$ and (S6) and (S7) both reduce to "$\mathfrak{M}, \ell \models \varphi$". So $\forall x\,\varphi$ and $\exists x\,\varphi$ are satisfied together or not at all, and every equation $x = y$ is satisfied. In any structure, $\forall x\,(x = x)$ is satisfied, since every d is the same element as itself; and $\exists x\,\neg(x = x)$ never is.

Satisfaction as a game. Section 5.1 read a sentence whose quantifiers all stand in front as a game between a challenger, who tries to make it false, and a defender, who tries to make it true. The clauses extend that game to every formula, one move per clause. At $\forall x\,\varphi$ (S6) the challenger picks an element $d \in D$ for x; at $\exists x\,\varphi$ (S7) the defender picks. At $\varphi \wedge \psi$ the challenger picks the side to be checked, and at $\varphi \vee \psi$ the defender does; $\varphi \to \psi$ is played as $\neg\varphi \vee \psi$, and $\varphi \leftrightarrow \psi$ as the two implications joined by $\wedge$. At $\neg\varphi$ the players change roles (S4), since making

$\neg\varphi$ true is making φ false. At an atomic formula the game ends, and the defender wins exactly when (S1) or (S2) holds; at $\bot$ the defender never wins (S3). A sentence is true in $\mathfrak{M}$ exactly when the defender can answer every challenge and win: read clause by clause, that is what (S1)–(S7) say. In the trace above, the defender's answer $e = d + 1$ to each challenge d is such a way of winning on $\mathbb{N}$; on $\{1, 2, 3\}$ the challenge $d = 3$ has no answer.

Quick check. Play the game for $\exists y \forall x\, (x \leq y)$ on $(\mathbb{N}, \leq)$ and on $(\{1, 2, 3\}, \leq)$. Who moves first, and who wins?

Solution. The defender moves first, since $\exists$ comes first: the order of the quantifiers is the order of the moves. On $\mathbb{N}$, whatever e the defender names, the challenger answers $x = e + 1$ and wins, so the sentence is false. On $\{1, 2, 3\}$ the defender names 3, and every challenge x has $x \leq 3$, so the sentence is true.

Quick check. Is $\forall x \exists y\, (y \cdot y = x)$ satisfied in $\mathbb{R}$, in $[0, \infty)$, and in $\mathbb{C}$, each with $\cdot$ the usual product?

Solution. In $\mathbb{R}$, no: at $d = -1$, (S7) asks for a real e with $e \cdot e = -1$, and every real square is ≥ 0. In $[0, \infty)$, yes: for each $d \geq 0$, $e = \sqrt{d}$ is in the domain and $e \cdot e = d$; and $[0, \infty)$ with $\cdot$ is a structure, since a product of non-negative reals is non-negative (M3). In $\mathbb{C}$, yes: every complex number has a complex square root (school algebra of complex numbers: $r(\cos\theta + i \sin\theta)$ with $r \geq 0$ has the square root $\sqrt{r}\,(\cos\frac{\theta}{2} + i \sin\frac{\theta}{2})$). The sentence is the same all three times; only the structure changes.

7.1.6 Only the free variables matter

The example of $\forall y\, (x \leq y)$ above showed ℓ mattering. In the first example it did not: the verdict on $\forall x \exists y\, (x < y)$ was reached for "any assignment ℓ". That is no accident. A sentence has no free variables (Definition 5.2.4), and every variable it contains is bound, so every value ℓ gives is overwritten by (S6) or (S7) before it is looked at. We prove the general fact, which is the first-order version of Chapter 4's Lemma 4.4.1: satisfaction depends only on the values of the **free** variables.

Say that two assignments ℓ and ℓ' **agree on** a set of variables if $\ell(y) = \ell'(y)$ for every variable y in it. We need the free variables of a formula (Definition 5.2.3), in the recursive description given after that definition in Section 5.2: every variable occurring in an atomic formula is free there; $\neg\varphi$ has the free variables of φ; $\varphi \circ \psi$, for a binary connective $\circ$, has those of φ and those of ψ; and $\forall x\, \varphi$ and $\exists x\, \varphi$ have those of φ other than x.

Lemma 7.1.1. Free Variables Coincidence Lemma

Let $\mathfrak{M}$ be a structure and φ a formula, and let ℓ and ℓ' be assignments in $\mathfrak{M}$ that agree on

every free variable of φ. Then

$$\mathfrak{M}, \ell \models \varphi \quad \text{if and only if} \quad \mathfrak{M}, \ell' \models \varphi.$$

In words: to decide a formula, only the values of its free variables matter.

Second reading · *The proof: an induction for all assignments at once*

Idea. Structural induction on φ, as for Lemma 4.4.1, after a first induction on terms: a term's value depends only on the values of its variables. The connective cases are Chapter 4's. The quantifier case is the only new one, and it explains why the property must be stated **for all** pairs of assignments: (S6) compares $\ell[x \mapsto d]$ with $\ell'[x \mapsto d]$, which are new assignments, so the induction hypothesis must apply to them. They agree at x, since both give it d, and at every other free variable of φ, since those are free in $\forall x\, \varphi$.

Proof. Terms first. We show: if ℓ and ℓ' agree on every variable occurring in a term t, then $t^{\mathfrak{M},\ell} = t^{\mathfrak{M},\ell'}$. We use structural induction on terms (Definition 5.2.1; the pattern of Theorem 4.3.1). For a variable x, the variable x occurs in x, so $x^{\mathfrak{M},\ell} = \ell(x) = \ell'(x) = x^{\mathfrak{M},\ell'}$. For a constant c, both values are $c^{\mathfrak{M}}$. For $f(t_1, \ldots, t_n)$, every variable of t_i occurs in $f(t_1, \ldots, t_n)$, so the hypothesis for t_i gives $t_i^{\mathfrak{M},\ell} = t_i^{\mathfrak{M},\ell'}$ for each i; applying $f^{\mathfrak{M}}$ to equal tuples gives equal values.

Formulas. We apply structural induction to the property $C(\varphi)$: *for all assignments ℓ, ℓ' in $\mathfrak{M}$ that agree on the free variables of φ, $\mathfrak{M}, \ell \models \varphi$ if and only if $\mathfrak{M}, \ell' \models \varphi$.*

Atomic formulas. Let φ be $R(t_1, \ldots, t_n)$, and let ℓ, ℓ' agree on its free variables. Every variable occurring in φ is free in it (Definition 5.2.3), so ℓ and ℓ' agree on the variables of each t_i, and by the claim for terms $t_i^{\mathfrak{M},\ell} = t_i^{\mathfrak{M},\ell'}$. The two tuples in (S1) are equal, so one lies in $R^{\mathfrak{M}}$ exactly when the other does. The case $t_1 = t_2$ is the same with (S2). For $\bot$, neither side holds, by (S3).

Connectives. Assume $C(\varphi)$ and $C(\psi)$, and let ℓ, ℓ' agree on the free variables of $\varphi \wedge \psi$. These include the free variables of φ and those of ψ, so $\mathfrak{M}, \ell \models \varphi$ if and only if $\mathfrak{M}, \ell' \models \varphi$, and likewise for ψ. By (S5), satisfaction of $\varphi \wedge \psi$ is decided by these two answers through the same rule for ℓ and for ℓ', so $C(\varphi \wedge \psi)$ holds. The cases $\neg\varphi$, $\varphi \vee \psi$, $\varphi \to \psi$, and $\varphi \leftrightarrow \psi$ are the same argument with (S4) and the other parts of (S5), as in the proof of Lemma 4.4.1.

Universal quantifier. Assume $C(\varphi)$, let x be a variable, and let ℓ, ℓ' agree on the free variables of $\forall x\, \varphi$. Let $d \in D$. We claim that $\ell[x \mapsto d]$ and $\ell'[x \mapsto d]$ agree on the free variables of φ. At x, both give d. A free variable y of φ other than x is a free variable of $\forall x\, \varphi$ (Definition 5.2.3), so $\ell[x \mapsto d](y) = \ell(y) = \ell'(y) = \ell'[x \mapsto d](y)$. By $C(\varphi)$, applied to these two assignments,

$$\mathfrak{M}, \ell[x \mapsto d] \models \varphi \quad \text{iff} \quad \mathfrak{M}, \ell'[x \mapsto d] \models \varphi.$$

This holds for every $d \in D$, so by (S6) $\mathfrak{M}, \ell \models \forall x\, \varphi$ if and only if $\mathfrak{M}, \ell' \models \forall x\, \varphi$. This proves $C(\forall x\, \varphi)$.

Existential quantifier. The same argument, with (S7) in place of (S6): the displayed equivalence holds for each d, so there is a d making the left side hold exactly when there is one

making the right side hold.

Every clause of Definition 5.2.2 has been covered, so $C(\varphi)$ holds for every formula φ. This proves the lemma.

■

Quick check. In $(\mathbb{N}, \leq)$, does whether ℓ satisfies $\forall y\,(x \leq y)$ depend on $\ell(y)$?

Solution. No: y is not free in the formula, so by the lemma only $\ell(x)$ matters. The formula holds exactly when $\ell(x) = 0$.

The lemma has a consequence that the rest of the section uses constantly.

Sentences. If φ is a sentence, it has no free variables, so **any** two assignments agree on its free variables, and the lemma says that $\mathfrak{M}, \ell \models \varphi$ either holds for every assignment ℓ or for none. We then write $\mathfrak{M} \models \varphi$, and say that φ is **true in** $\mathfrak{M}$ (or that $\mathfrak{M}$ is a model of φ), in the first case, and $\mathfrak{M} \not\models \varphi$, φ is **false in** $\mathfrak{M}$, in the second. At least one assignment exists because D is non-empty, so a sentence is always exactly one of true and false in $\mathfrak{M}$. This is the precise form of Section 5.1's informal "true over a domain".

Only finitely many values matter. A formula has finitely many free variables. To decide $\mathfrak{M}, \ell \models \varphi$, it is enough to know ℓ on those. So we describe an assignment by the values that matter ("$\ell(x) = 2$"), exactly as Chapter 4 described a valuation by the values of a few atoms.

Warning. The same sentence can be true in one structure and false in another, and "true" without a structure means nothing. $\forall x\,\exists y\,(x < y)$ is true in $(\mathbb{N}, <)$ and false in $(\{1, 2, 3\}, <)$, as traced above. A sentence of predicate logic is like a formula of Chapter 4 with atoms: it has a truth value only once the meaning of its symbols has been fixed.

7.1.7 Semantic entailment

Chapter 4 said that Γ entails ψ when every row that makes Γ true makes ψ true (Definition 4.5.3). The row is now a structure with an assignment.

An argument is valid when every structure, under every assignment, that satisfies all the premises also satisfies the conclusion.

Definition 7.1.4. Semantic Entailment in Predicate Logic

Let Γ be a list of formulas and φ a formula. We say Γ **semantically entails** φ, and write $\Gamma \models \varphi$, if for **every** structure $\mathfrak{M}$ for the symbols occurring in Γ and φ, and **every** assignment ℓ in $\mathfrak{M}$,

$$\mathfrak{M}, \ell \models \Gamma \quad \Rightarrow \quad \mathfrak{M}, \ell \models \varphi.$$

We write $\models \varphi$ when Γ is empty. If $\Gamma \models \varphi$ fails, we write $\Gamma \not\models \varphi$; a structure $\mathfrak{M}$ with an

assignment ℓ such that $\mathfrak{M}, \ell \models \Gamma$ and $\mathfrak{M}, \ell \not\models \varphi$ is a **counter-model** to $\Gamma \models \varphi$.

In words: "every structure" means every domain and every meaning of the symbols, so an entailment holds for reasons of form and not because of what R happens to mean. "Every assignment" covers every value of the free variables; when Γ and φ are sentences, the assignment makes no difference, by Lemma 7.1.1. And a counter-model is the analogue of Chapter 4's refuting row: one structure, with one assignment, where the premises hold and the conclusion does not. The symbol $\models$ is the one of Chapter 4, and the context always shows which is meant: a sequent of formulas with quantifiers or predicate symbols is read with this definition.

A structure may interpret symbols that do not occur in Γ or φ. The clauses of Definition 7.1.3 consult only the symbols of the formula at hand, so such extra meanings change nothing.

First examples of entailment. The argument Section 5.1 opened with, "every positive real has a square root; 2 is positive; so 2 has a square root", is $\forall x\,(P(x) \to Q(x)),\ P(c) \vdash Q(c)$, derived in Section 5.3. Truth tables called it invalid, since they saw three unrelated atoms. It is valid in the new sense:

$$\forall x\,(P(x) \to Q(x)),\ P(c) \models Q(c).$$

Let $\mathfrak{M}, \ell$ satisfy both premises, and let $a = c^{\mathfrak{M}}$. By (S6) for the first premise, with $d = a$, we get $\mathfrak{M}, \ell[x \mapsto a] \models P(x) \to Q(x)$. Under $\ell[x \mapsto a]$ the term x has value a, and under ℓ the term c has value a; so by (S1), $\mathfrak{M}, \ell[x \mapsto a] \models P(x)$ holds if and only if $a \in P^{\mathfrak{M}}$ if and only if $\mathfrak{M}, \ell \models P(c)$, which is the second premise. So by (S5) $\mathfrak{M}, \ell[x \mapsto a] \models Q(x)$, that is, $a \in Q^{\mathfrak{M}}$, and by (S1) again $\mathfrak{M}, \ell \models Q(c)$. Chapter 4's objection to this argument is gone.

7.1.8 Proving an entailment by meaning

To show $\Gamma \models \varphi$, fix an arbitrary structure and assignment satisfying Γ, and show that they satisfy φ. The first proof below writes every step with the clause that licenses it; the second condenses it to what a mathematician writes.

Example. A Semantic Proof, Clause by Clause

Show that $\exists x\,\forall y\,R(x, y) \models \forall y\,\exists x\,R(x, y)$.

Solution. Each step ends with the clause that licenses it.

1. Let $\mathfrak{M}$ be a structure with domain D and ℓ an assignment with $\mathfrak{M}, \ell \models \exists x\,\forall y\,R(x, y)$ (Definition 7.1.4).
2. There is $d \in D$ with $\mathfrak{M}, \ell[x \mapsto d] \models \forall y\,R(x, y)$ (S7).
3. To show $\mathfrak{M}, \ell \models \forall y\,\exists x\,R(x, y)$, let $e \in D$; we must show $\mathfrak{M}, \ell[y \mapsto e] \models \exists x\,R(x, y)$ (S6).
4. From item 2, with this e: $\mathfrak{M}, \ell[x \mapsto d][y \mapsto e] \models R(x, y)$ (S6).
5. Under $\ell[x \mapsto d][y \mapsto e]$ the terms x, y have the values d, e, so $(d, e) \in R^{\mathfrak{M}}$ (S1).
6. Take the witness d for x: under $\ell[y \mapsto e][x \mapsto d]$ the values are again d, e, so $\mathfrak{M}, \ell[y \mapsto e][x \mapsto d] \models R(x, y)$ by item 5 (S1), and so $\mathfrak{M}, \ell[y \mapsto e] \models \exists x\,R(x, y)$ (S7).
7. Since e was arbitrary, $\mathfrak{M}, \ell \models \forall y\,\exists x\,R(x, y)$ (S6).

Condensed. Suppose some d is R-related to everything. Then for each e, the element d itself is related to e, so each e has something related to it.

Section 5.6 derived the same sequent (Proposition 5.6.2); this is its meaning. The converse fails, and Section 7.2 draws why.

Quick check. In item 6, why are $\ell[y \mapsto e][x \mapsto d]$ and $\ell[x \mapsto d][y \mapsto e]$ the same assignment?

Solution. The variables x and y are different, and each change touches only its own variable, so the order of the two changes does not matter (the paragraph after Definition 7.1.2).

Example. The Drinker Sentence, by Meaning

Show that $\models \exists x\,(P(x) \to \forall y\,P(y))$: in every structure, some element has the property that if **it** has P, everything has P.

Solution. Let $\mathfrak{M}$ and ℓ be given, with domain D.

Case 1: $P^{\mathfrak{M}} = D$. Choose any $d \in D$, which exists because D is non-empty (M1). Under $\ell[x \mapsto d]$ the formula $\forall y\,P(y)$ is satisfied, by (S6), so the implication is satisfied, by (S5); hence the sentence holds, by (S7).

Case 2: some $d \in D$ is not in $P^{\mathfrak{M}}$. Under $\ell[x \mapsto d]$ the hypothesis $P(x)$ fails (S1), so the implication is satisfied (S5), and (S7) gives the sentence.

The case split is on a question about the whole structure, whether $P^{\mathfrak{M}} = D$, and the witness depends on its answer; this is why no witness can be named in advance, and why a derivation needs a detour (Section 6.3 derived it, with D, n, m in place of P, x, y, Exercise (C1: Some index decides, used later)). Case 1 is where the non-empty domain is used, as in the next subsection.

7.1.9 Why domains are non-empty

Clause (M1) of Definition 7.1.1 demands a non-empty domain. Here is the reason. Proposition 5.5.1 derives $\forall x\,P(x) \vdash \exists x\,P(x)$ in three lines: $\forall$E at the variable x, then $\exists$I. The matching entailment holds for our structures:

$$\forall x\,P(x) \models \exists x\,P(x).$$

Let $\mathfrak{M}, \ell \models \forall x\,P(x)$. Since D is non-empty, choose some $d \in D$. By (S6), $\mathfrak{M}, \ell[x \mapsto d] \models P(x)$, and this d is the one element (S7) asks for, so $\mathfrak{M}, \ell \models \exists x\,P(x)$.

Now suppose empty domains were allowed. In a structure with $D = \emptyset$, (S6) says $\forall x\,P(x)$ is satisfied, since "for every $d \in \emptyset$" has nothing to check; and (S7) says $\exists x\,P(x)$ is not, since there is no d at all. So the sequent would be derivable and yet fail in a structure, and the rules would lead from a true premise to a false conclusion. (Worse, no assignment into an empty domain exists, so even the clauses could not be applied as written.) One of two things had to give: the rule $\forall$E at a variable, or empty domains. The book, like most of mathematics, keeps the rules and makes domains non-empty.

Nothing is lost. Statements about an empty **set** are handled by the bounded quantifiers of Definition 5.1.1, inside a non-empty domain. The bounded forms $\forall x \in \emptyset\ \varphi$ and $\exists x \in \emptyset\ \varphi$ abbreviate $\forall x\,(x \in \emptyset \to \varphi)$ and $\exists x\,(x \in \emptyset \wedge \varphi)$, and in a structure with domain $\mathbb{R}$ in which the one-place predicate symbol $\in \emptyset$ (Section 5.2) is interpreted, by (M4), as the empty subset of $\mathbb{R}$,

the first is true (every hypothesis $d \in \emptyset$ is false) and the second is false (no d is in $\emptyset$). Vacuous truth lives in the bound, not in the domain.

7.1.10 Exercises

7.1.10.1 A. Check your understanding

Exercise. A1

Let $\mathfrak{M}_{\mathbb{N}}$, $\mathfrak{M}_{\mathbb{Z}}$, $\mathfrak{M}_{\mathbb{Q}}$ be the structures with domains $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, each with $<$ the usual order.

(a) For each of the sentences (i) $\forall x\, \exists y\, (x < y)$, (ii) $\exists x\, \forall y\, \neg(y < x)$, (iii) $\forall x\, \forall y\, (x < y \to \exists z\, (x < z \wedge z < y))$ and (iv) $\forall x\, \exists y\, (y < x)$, say in which of the three structures it is true, with a witness or a failing element for each verdict.

(b) True or false: if a sentence is true in some structure, it is true in every structure for its symbols.

(c) Why must the domain of a structure be non-empty? Name the derivable sequent that would otherwise fail.

(d) In Example (A Semantic Proof, Clause by Clause), which clause produces the element d, and why may the same d serve for every e?

Solution. (a) (i) True in all three: for each d, take $y = d + 1$, which lies in the same domain and is larger.

(ii) This says "there is a least element". True in $\mathfrak{M}_{\mathbb{N}}$, with $x = 0$: no natural number is below 0. False in $\mathfrak{M}_{\mathbb{Z}}$ and $\mathfrak{M}_{\mathbb{Q}}$: for any d, the element $e = d - 1$ has $e < d$, so (S6) fails for $\neg(y < x)$ under $\ell[x \mapsto d]$.

(iii) This says "between any two elements there is a third" (density). False in $\mathfrak{M}_{\mathbb{N}}$ and $\mathfrak{M}_{\mathbb{Z}}$: at $x = 1$, $y = 2$ there is no integer strictly between. True in $\mathfrak{M}_{\mathbb{Q}}$: for $d < e$ in $\mathbb{Q}$, $z = (d + e)/2$ is rational and $d < z < e$.

(iv) This says "no element is least". False in $\mathfrak{M}_{\mathbb{N}}$, at $x = 0$. True in $\mathfrak{M}_{\mathbb{Z}}$ and $\mathfrak{M}_{\mathbb{Q}}$, with $y = d - 1$.

	$\mathbb{N}$	$\mathbb{Z}$	$\mathbb{Q}$
(i)	true	true	true
(ii)	true	false	false
(iii)	false	false	true
(iv)	false	true	true

(b) False. Sentence (ii) is true in $\mathfrak{M}_{\mathbb{N}}$ and false in $\mathfrak{M}_{\mathbb{Z}}$, as the table shows.

(c) The sequent is $\forall x\, P(x) \vdash \exists x\, P(x)$, derived in Proposition 5.5.1 by $\forall$E at the variable x and $\exists$I. In an empty domain $\forall x\, P(x)$ would be true, with nothing to check, and $\exists x\, P(x)$ false, with no witness; so the rules would lead from a true premise to a false conclusion. (There would also be no assignments at all.)

(d) (S7), applied to the premise in item 2. The element d is fixed before any e is chosen, and item 2 says that d is related to **every** e, by (S6); so item 6 may use d whatever e item 3 picked.

Exercise. A2: Drill

Each part takes a line or two. Parts (a)–(d) read a structure or a clause, (e)–(i) decide a formula in a given structure, with a witness or a failing element, (j) and (k) judge, and (l) builds a structure.

(a) Let $D = \{a, b\}$ and $P^{\mathfrak{M}} = \{a\}$. Is $\forall x\, P(x)$ true in $\mathfrak{M}$?

(b) In the same structure, is $\exists x\, \neg P(x)$ true?

(c) In $(\mathbb{R}; 1, +, \cdot)$, what is the value of $x \cdot x + 1$ under an assignment with $\ell(x) = -2$?

(d) By Lemma 7.1.1, the values of which variables **can** matter to the truth of $\forall y\, (x \leq y) \wedge z = z$ in a given structure? Do they all matter?

(e) In $(\{0, 1, 2\}, \leq)$, is $\exists x\, \forall y\, (x \leq y)$ true?

(f) In $(\mathbb{N}, <)$, under an assignment with $\ell(x) = 3$, is $\exists y\, (y < x)$ satisfied? And with $\ell(x) = 0$?

(g) In the left structure of Figure F, is $\forall x\, (P(x) \to \exists y\, R(y, x))$ true?

(h) Decide $\forall x\, \forall y\, (x \leq y \vee y \leq x)$ in $(\mathbb{N}, \leq)$ and in the right structure of Figure F.

(i) Decide $\exists x\, \forall y\, (y \cdot x = y)$ in $(\mathbb{Z}, \cdot)$ and in the even integers with $\cdot$.

(j) Is $\mathbb{N}$ with the two-place function symbol $-$ read as subtraction a structure?

(k) In a structure with $D = \{a\}$ and $P^{\mathfrak{M}} = \emptyset$, decide $\forall x\, P(x) \to \exists x\, P(x)$ and $\forall x\, \neg P(x) \to \neg \exists x\, P(x)$.

(l) Give a structure on $\{1, 2\}$ in which $\exists x\, R(x, x)$ is false and $\forall x\, \exists y\, R(x, y)$ is true.

Solution.

(a) No: it fails at b.

(b) Yes, at b.

(c) $(-2) \cdot (-2) + 1 = 5$.

(d) At most x and z, its free variables; not y. Not always: z never matters, since $z = z$ holds under every assignment, and whether x matters depends on the structure. In $(\mathbb{N}, \leq)$ it does (true with $\ell(x) = 0$, false with $\ell(x) = 1$); if $\leq^{\mathfrak{M}}$ is empty it does not. The lemma gives an upper bound, not the exact set.

(e) Yes, with $x = 0$.

(f) Yes, with $y = 0$. With $\ell(x) = 0$, no: nothing in $\mathbb{N}$ is below 0.

(g) Yes: the only element with P is a, and the arrow $c \to a$ enters it; at b and c the hypothesis fails.

(h) True in $\mathbb{N}$; false in the divisibility structure, at $x = 2$, $y = 3$.

(i) True in $\mathbb{Z}$, with $x = 1$. False in the even integers: $y = 2$ would need $2x = 2$, so $x = 1$, which is odd.

(j) No: (M3) fails, since $0 - 1 \notin \mathbb{N}$.

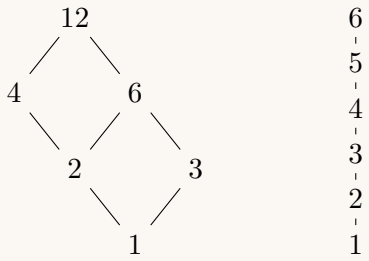
(k) Both true: the first because its hypothesis fails at a; the second because both sides are true.

(l) $R^{\mathfrak{M}} = \{(1, 2), (2, 1)\}$: no loops, and each element has an arrow out.

7.1.10.2 B. Practice

Exercise. B1: One sentence, two pictures

Two structures for the one symbol $\leq$ are drawn below. On the left, $\mathfrak{M}_1$: the domain $D_1 = \{1, 2, 3, 4, 6, 12\}$, with $x \leq y$ read as "x divides y", as a Hasse diagram. On the right, $\mathfrak{M}_2 = (\{1, 2, 3, 4, 5, 6\}, \leq)$, with the usual order, as a vertical chain.



Decide each sentence in each structure, with a witness or a failing pair:

(i) $\exists x \, \forall y \, (x \leq y)$;

(ii) $\exists x \, \forall y \, (y \leq x)$;

(iii) $\forall x \, \forall y \, (x \leq y \vee y \leq x)$;

(iv) $\forall x \, \forall y \, \exists z \, (x \leq z \wedge y \leq z)$.

Which of them tell the two structures apart?

Solution. *(i)* True in both: $x = 1$ in each, since 1 divides every element of D_1 and is the least element of $\mathfrak{M}_2$.

(ii) True in both: $x = 12$ in $\mathfrak{M}_1$, since every element of D_1 divides 12; $x = 6$ in $\mathfrak{M}_2$.

(iii) False in $\mathfrak{M}_1$, at $x = 2$, $y = 3$: neither divides the other. True in $\mathfrak{M}_2$: of two numbers, one is at most the other.

(iv) True in both: $z = 12$ in $\mathfrak{M}_1$, and $z = 6$ in $\mathfrak{M}_2$ (or z the larger of x and y).

Only (iii) separates them.

Marking scheme. (10 points) - 8: 2 for each of (i)–(iv) in both structures, 1 for the verdicts and 1 for a named witness or failing pair for every verdict. - 2: naming (iii) as the only separating sentence. Fatal errors (0 for the part): a verdict on $\mathfrak{M}_1$ that reads $\leq$ as the usual order of numbers.

Exercise. B2: A model on three elements

Give an example of a structure with exactly three elements in which both

$$\forall x\, \exists y\, R(x,y) \qquad \text{and} \qquad \forall x\, \neg R(x,x)$$

are true. Check both sentences. Which properties must every such structure have?

Solution. Take $D = \{0,1,2\}$ and $R^{\mathfrak{M}} = \{(0,1),(1,2),(2,0)\}$: each element is related to the next one around a circle. First sentence: for $d = 0,1,2$, the witnesses $e = 1,2,0$ give pairs in $R^{\mathfrak{M}}$. Second sentence: none of $(0,0)$, $(1,1)$, $(2,2)$ is in $R^{\mathfrak{M}}$, so $\neg R(x,x)$ is satisfied under $\ell[x \mapsto d]$ for each d.

The properties every answer must have: the domain has three elements; every element d has **some** e with $(d,e) \in R^{\mathfrak{M}}$; and **no** pair (d,d) is in $R^{\mathfrak{M}}$, so that e is always different from d. Any $R^{\mathfrak{M}}$ meeting these works, for instance $\{(0,1),(1,0),(2,0)\}$ or the set of all pairs (d,e) with $d \neq e$.

Exercise. B3: Is it a structure?

For each description, say whether it is a structure for the symbols named. If not, name the clause of Definition 7.1.1 that fails, and repair it with as small a change as you can.

(i) $D = \mathbb{Q}$, with a one-place function symbol f read as $x \mapsto 1/x$ for $x \neq 0$ and $f(0) = 0$.
(ii) $D = \mathbb{N}$, with f read as $x \mapsto x - 1$.
(iii) $D = \emptyset$, with P read as the empty set.
(iv) $D = \mathbb{Z}$, with a constant c naming $\frac{1}{2}$.
(v) $D = \mathbb{R}$, with a two-place predicate symbol R read as $\{(x,y) : y = \sqrt{x}\}$.
(vi) $D = \mathbb{R}$, with a one-place function symbol g read as $x \mapsto \sqrt{x}$.

Solution. *(i)* Yes: every element has a value in $\mathbb{Q}$.

(ii) No: (M3) fails, since $f(0) = -1 \notin \mathbb{N}$. Repair: $f(0) \coloneqq 0$.

(iii) No: (M1) fails. Repair: any non-empty domain.

(iv) No: (M2) fails, since $\frac{1}{2} \notin \mathbb{Z}$. Repair: $D = \mathbb{Q}$, or let c name an integer.

(v) Yes: (M4) asks only for a set of pairs, and nothing requires a relation to hold for every first element.

(vi) No: (M3) fails, since $\sqrt{-1}$ is not real. Repair: $D = [0,\infty)$, or fix $g(x) \coloneqq 0$ for $x < 0$.

Descriptions (v) and (vi) carry the same information, and only the function symbol must be total.

Marking scheme. (10 points) - 6: 1 for each verdict. - 2: the clause named at each "no" (all four right). - 2: repairs that are minimal and correct. Fatal errors (0 for the part): calling (v) a non-structure because $\sqrt{x}$ is undefined for $x < 0$.

7.1.10.3 C. Going deeper

Exercise. C1: Proofs by meaning

(a) Prove $\forall x\,(P(x) \wedge Q(x)) \models \forall x\,P(x) \wedge \forall x\,Q(x)$ clause by clause, as in Example (A Semantic Proof, Clause by Clause).

(b) Prove $\forall x\,(P(x) \vee Q(x)) \models \forall x\,P(x) \vee \exists x\,Q(x)$, condensed.

(c) Prove $\exists x\,(P(x) \to Q(c)) \models \forall x\,P(x) \to Q(c)$, condensed, and say where your proof uses the existential premise.

Solution. *(a)* Let $\mathfrak{M}, \ell \models \forall x\,(P(x) \wedge Q(x))$. By (S5) it suffices to show each conjunct. For $\forall x\,P(x)$, let $d \in D$ (S6). By (S6) for the premise, $\mathfrak{M}, \ell[x \mapsto d] \models P(x) \wedge Q(x)$; by (S5), $\mathfrak{M}, \ell[x \mapsto d] \models P(x)$. Since d was arbitrary, $\mathfrak{M}, \ell \models \forall x\,P(x)$ (S6). Likewise $\mathfrak{M}, \ell \models \forall x\,Q(x)$, with the other half of (S5).

(b) If $P^{\mathfrak{M}} = D$, the left disjunct holds. Otherwise some $d \in D$ is not in $P^{\mathfrak{M}}$. The premise at d gives $d \in P^{\mathfrak{M}}$ or $d \in Q^{\mathfrak{M}}$, so $d \in Q^{\mathfrak{M}}$, and $\exists x\,Q(x)$ holds. Either way the disjunction holds.

(c) Let d be the premise's witness: if $d \in P^{\mathfrak{M}}$, then $c^{\mathfrak{M}} \in Q^{\mathfrak{M}}$. Suppose $\forall x\,P(x)$ holds; then $d \in P^{\mathfrak{M}}$, so $Q(c)$ holds. The witness d is where the existential premise is used: the hypothesis $\forall x\,P(x)$ is applied at it.

Marking scheme. (10 points) - 4: (a), with every step's clause named, and (S6) used both to fix d and on the premise. - 3: (b), split on whether $P^{\mathfrak{M}} = D$ (or on a failing element), with the premise used at that element. - 3: (c), taking the premise's witness d and applying the hypothesis $\forall x\,P(x)$ at d (in either order). Fatal errors (0 for the part): in (a), a proof in which d is a particular element (for example a constant's value, or an element chosen to lie in $P^{\mathfrak{M}}$) rather than an arbitrary $d \in D$, so that the closing (S6) step for $\forall x\,P(x)$ is unjustified.

Exercise. C2: Challenge: only the symbols that occur matter

Let $\mathfrak{M}$ and $\mathfrak{N}$ be structures with the same domain D that give the same meaning to every symbol occurring in a formula φ. Prove that for every assignment ℓ, $\mathfrak{M}, \ell \models \varphi$ if and only if $\mathfrak{N}, \ell \models \varphi$. (This is the claim after Definition 7.1.4 that extra meanings change nothing.)

Solution (Hint 1). Terms first, as in the proof of Lemma 7.1.1.

Solution (Hint 2). Prove, by structural induction on φ, the property: for **every** assignment ℓ, ...; the quantifier case needs it at $\ell[x \mapsto d]$.

Solution. *Terms.* For every term t whose symbols $\mathfrak{M}$ and $\mathfrak{N}$ interpret alike, $t^{\mathfrak{M},\ell} = t^{\mathfrak{N},\ell}$ for every ℓ, by structural induction on t. For a variable x, both values are $\ell(x)$. For a constant c, they are $c^{\mathfrak{M}} = c^{\mathfrak{N}}$. For $f(t_1, \ldots, t_n)$, the hypothesis gives equal values of the t_i, and $f^{\mathfrak{M}} = f^{\mathfrak{N}}$ applied to equal tuples gives equal values.

Formulas. Let $A(\varphi)$ be the property: *for every assignment* ℓ, $\mathfrak{M}, \ell \models \varphi$ *if and only if* $\mathfrak{N}, \ell \models \varphi$, for formulas φ whose symbols $\mathfrak{M}$ and $\mathfrak{N}$ interpret alike. Atomic formulas: the tuples of values are equal by the claim for terms, and $R^{\mathfrak{M}} = R^{\mathfrak{N}}$, so (S1) gives the same answer; equations likewise by (S2); $\bot$ by (S3). Connectives: (S4) and (S5) combine equal answers into equal answers. For $\forall x\, \varphi$: the two structures have the same domain, so d ranges over the same elements in both; for each d, $A(\varphi)$ at the assignment $\ell[x \mapsto d]$ gives equal answers, and (S6) gives $A(\forall x\, \varphi)$. For $\exists x\, \varphi$, the same with (S7). Every clause of Definition 5.2.2 has been covered, so $A(\varphi)$ holds for every formula.

Marking scheme. (10 points) - 2: the claim for terms, by induction. - 2: the atomic cases, with (S1) and (S2). - 2: the connective cases. - 3: the quantifier case, using the property at the changed assignments and the shared domain. - 1: naming where "same domain" is used. Fatal errors (0 for the step): an induction hypothesis stated for one fixed ℓ only.

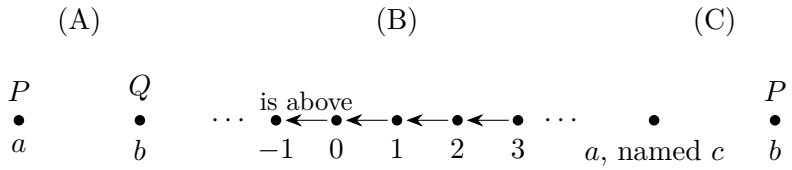
7.2 Counter-models

Section 7.1 gave every formula a meaning in a structure. This section uses it to answer the question Chapters 5 and 6 left open: when a search for a derivation stalls, is there no derivation, or only no derivation found yet? The answer is a **counter-model**, a structure with an assignment in which the premises are true and the conclusion false (Definition 7.1.4), together with one statement, soundness for predicate logic, that turns it into a proof that no derivation exists.

Remark (Looking back). Sections 4.5 and 4.6 refuted a propositional claim by a row that made its counterexample set true, and searched for that row systematically; Section 4.7 proved that a refuting row rules out every derivation. This section does the same with structures: the counter-model is the row, the procedure below is the search, and the side-condition table plays the part of Section 4.7's rule-by-rule table.

7.2.1 Three pictures

Section 5.3 called a reading in which the premises are true and the conclusion false a counter-reading. Here are three, drawn as in Section 7.1.



Three counter-readings. (A) P *holds at* a *and* Q *at* b: $\exists x\, P(x) \wedge \exists x\, Q(x)$ *is true, with the witnesses* a *and* b, *and* $\exists x\, (P(x) \wedge Q(x))$ *is false, since no dot has both labels; so the converse of Proposition 5.6.1 does not follow (Section 5.6 asked). (B) The integers, with* $R(x, y)$ *read as* $x > y$. *Only the arrows to the next integer down are drawn;* R *is the whole order* $>$, *so* $R(x, y)$ *holds whenever a chain of arrows leads from* x *to* y. $\forall y\, \exists x\, R(x, y)$ *is true, since every integer* y *has* $y + 1$ *above it, and* $\exists x\, \forall y\, R(x, y)$ *is false, since an integer above every integer would be above itself; so the converse of Proposition 5.6.2 does not follow. (C)* P *holds at* b, *and the constant* c *names* a: $\exists x\, P(x)$ *is true, at* b, *and* $P(c)$ *is false (Exercise (C2: A witness is not a name, used later) asked why no derivation was found).*

Each picture is a structure in the sense of Section 7.1: a domain with a meaning for every symbol. The rest of this section checks such pictures by the clauses, finds them by a procedure, and shows why one of them rules out every derivation.

Quick check. Draw two dots that make $\forall x\,(P(x) \lor Q(x))$ true and $\forall x\,P(x) \lor \forall x\,Q(x)$ false.

Solution. Picture (A) again: each dot has P or Q; $\forall x\,P(x)$ fails at b, and $\forall x\,Q(x)$ fails at a.

Quick check. In picture (C), which clauses of Definition 7.1.3 decide $\exists x\,P(x)$, and at which element?

Solution. (S7) asks for some $d \in D$ with $\mathfrak{M}, \ell[x \mapsto d] \models P(x)$; at $d = b$, (S1) asks whether $b \in P^{\mathfrak{M}}$, and it is.

7.2.2 Counter-models

A counter-model is found the way a refuting row was found in Chapter 4, by working backwards from what must fail. The conclusion must be false, so read off what that demands; the premises must be true, so read off what they demand; and then build the smallest structure meeting both. The worked examples that follow (the swap, the converses of Proposition 5.6.1 and of the $\forall/\lor$ law, $\exists x\,P(x) \vdash P(c)$, and $P(x_0) \vdash \forall x\,P(x)$) are sequents whose derivations stalled or were blocked in earlier sections.

Example. The swap of $\exists$ and $\forall$

Show that

$$\forall y\,\exists x\,R(x,y) \not\models \exists x\,\forall y\,R(x,y),$$

first by a structure on $\mathbb{Z}$, then by a structure with two elements.

Solution. Both formulas are sentences, so by Lemma 7.1.1 we need only a structure; any assignment will do.

A structure on $\mathbb{Z}$. Let $\mathfrak{M}$ have domain $\mathbb{Z}$ and $R^{\mathfrak{M}} = \{(x,y) : x > y\}$, so that $R(x,y)$ says "$x > y$". The premise says "every integer has an integer above it". To check it by (S6) and (S7): for each $e \in \mathbb{Z}$ we need some $d \in \mathbb{Z}$ with $(d,e) \in R^{\mathfrak{M}}$, that is, $d > e$; $d = e + 1$ works. The conclusion says "some integer is above every integer". By (S7) and (S6) it would need a $d \in \mathbb{Z}$ with $d > e$ for **every** $e \in \mathbb{Z}$; but $e = d$ gives $d > d$, which is false. So $\mathfrak{M} \models \forall y\,\exists x\,R(x,y)$ and $\mathfrak{M} \not\models \exists x\,\forall y\,R(x,y)$.

A structure with two elements. Let $D = \{a,b\}$ and $R^{\mathfrak{M}} = \{(a,a),(b,b)\}$, so that $R(x,y)$ says "x is y". Premise: for $e = a$ take $d = a$, and for $e = b$ take $d = b$; both pairs are in $R^{\mathfrak{M}}$. Conclusion: $d = a$ fails at $e = b$, since $(a,b) \notin R^{\mathfrak{M}}$, and $d = b$ fails at $e = a$. So the premise is true and the conclusion false.

The pattern is the one Section 5.1 read as a game. In the premise the witness x is chosen **after** y and may depend on it ($x = y + 1$, or $x = y$); the conclusion asks for one x that works for every y in advance. A one-element structure cannot separate the two sentences: with $D = \{a\}$, both are true if $(a, a) \in R^{\mathfrak{M}}$ and both false otherwise. Two elements are the least that will do.

The two-element structure is the $\mathbb{Z}$ structure's idea at its smallest: each element is related to something, and nothing is related to everything.

Picture (A), checked by the clauses. With $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$, $Q^{\mathfrak{M}} = \{b\}$, the premise of the converse of Proposition 5.6.1 holds by (S5) and (S7), with witnesses a and b, and the conclusion fails, since $a \notin Q^{\mathfrak{M}}$ and $b \notin P^{\mathfrak{M}}$:

$$\exists x\, P(x) \land \exists x\, Q(x) \not\models \exists x\, (P(x) \land Q(x)).$$

The same structure refutes the converse of the $\forall/\lor$ law of Proposition 6.2.6, since each element is in $P^{\mathfrak{M}}$ or in $Q^{\mathfrak{M}}$, and neither set is all of D:

$$\forall x\, (P(x) \lor Q(x)) \not\models \forall x\, P(x) \lor \forall x\, Q(x).$$

Picture (C): with $c^{\mathfrak{M}} = a$ and $P^{\mathfrak{M}} = \{b\}$, $\exists x\, P(x)$ holds with the witness b, and $P(c)$ fails, since $a \notin P^{\mathfrak{M}}$. So $\exists x\, P(x) \not\models P(c)$. In $\mathbb{Z}$, with P and Q read as "even" and "odd": some integer is even and some integer is odd, but no integer is both; and every integer is even or odd, but it is false that every integer is even, and false that every integer is odd.

The assignment is part of a counter-model. Section 5.4 showed why $\forall$I may not generalize over a name in a premise: from $P(x_0)$ one would "derive" $\forall x\, P(x)$. Here x_0 is a **variable**, free in the premise, so a counter-model needs an assignment as well as a structure. Let $\mathfrak{M}$ have domain $\mathbb{Z}$ with $P^{\mathfrak{M}}$ the set of even integers, and let $\ell(x_0) = 2$. Then $\mathfrak{M}, \ell \models P(x_0)$, since 2 is even, and $\mathfrak{M}, \ell \not\models \forall x\, P(x)$, since (S6) fails at $d = 1$. So $P(x_0) \not\models \forall x\, P(x)$. Under the assignment with $\ell(x_0) = 1$ the premise is false, and that assignment refutes nothing; the counter-model is the structure **with** $\ell(x_0) = 2$.

Warning. A "counter-model" that falsifies a premise refutes nothing. For $\forall y\, \exists x\, R(x, y) \vdash \exists x\, \forall y\, R(x, y)$, the structure $(\mathbb{N}, <)$, with $R(x, y)$ read as "$x < y$", makes the conclusion false, since no natural number is below every natural number; but it also makes the premise false, since nothing in $\mathbb{N}$ is below $y = 0$. It is not a counter-model, and it proves nothing about the sequent. As with rows in Chapter 4, check the premises **first**.

A model of a list of sentences. Counter-models are one use of structures; describing a whole kind of object is another. Consider the three sentences

$$\forall x\, \neg(x < x),$$
$$\forall x\, \forall y\, \forall z\, ((x < y \land y < z) \to x < z),$$
$$\forall x\, \exists y\, (x < y),$$

which say that $<$ is a **strict order** (never $x < x$; $x < y$ and $y < z$ give $x < z$) **with no largest element**. The structure $(\mathbb{N}, <)$ is a model of all three: no number is less than itself, the order is transitive, and $d + 1 > d$ is a witness for the third sentence. It is the one the sentences were written with in mind, but not the only one: in $(\mathbb{Z}, <)$ and $(\mathbb{Q}, <)$ all three are true, by the same checks.

Second reading · *Finite and infinite models*

Finite counter-models are convenient because every clause can be checked by listing. They are not always available: Exercise C1 below gives a sentence that is true in every finite structure and false in an infinite one, and Exercise C2 a sentence with only infinite models.

These models differ in ways the three sentences cannot see: $\mathbb{Z}$ has no least element, and $\mathbb{Q}$ has a third element between any two. And the one-element structure $D = \{a\}$ is not a model of the three sentences: if $(a, a) \in <^{\mathfrak{M}}$ the first sentence fails, and if not, the third fails at $d = a$. Exercise C2 shows that no **finite** structure is a model of them.

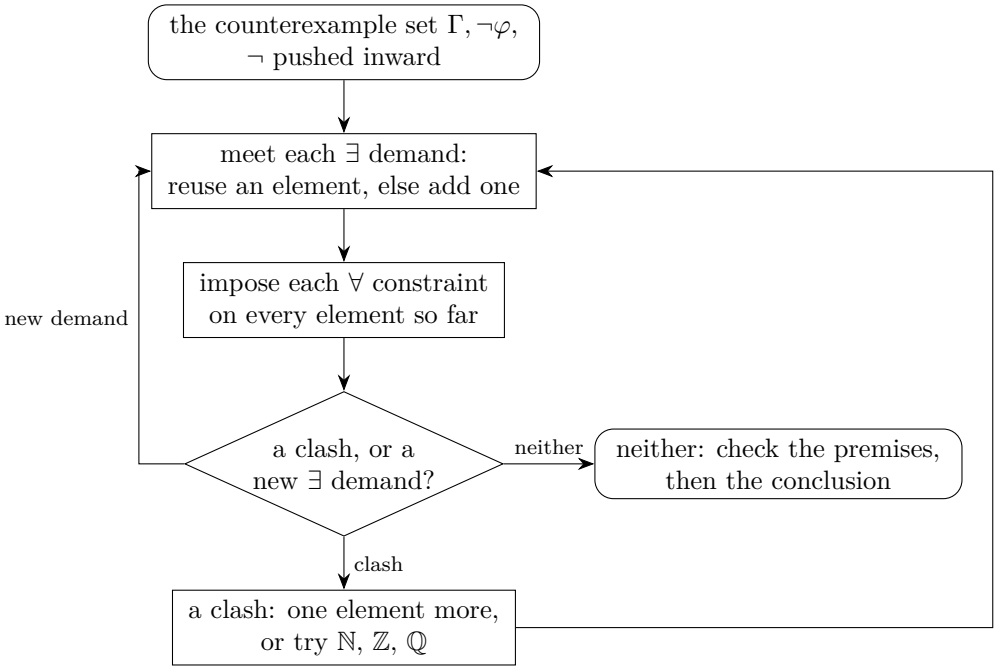
Quick check. Give a counter-model to $\forall x\, P(x) \to \forall x\, Q(x) \models \forall x\, (P(x) \to Q(x))$. (Think about what makes an implication with a universal hypothesis true cheaply.)

Solution. Take $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$, $Q^{\mathfrak{M}} = \emptyset$. The premise is an implication whose hypothesis $\forall x\, P(x)$ is false (it fails at b), so by (S5) the premise is true. The conclusion fails at $d = a$: $a \in P^{\mathfrak{M}}$ and $a \notin Q^{\mathfrak{M}}$, so $P(x) \to Q(x)$ is not satisfied under $\ell[x \mapsto a]$. (Any structure with $P^{\mathfrak{M}} \neq D$ and some element in $P^{\mathfrak{M}}$ but not in $Q^{\mathfrak{M}}$ works.)

7.2.3 Finding a counter-model

A counter-model is a structure with an assignment under which every member of the **counterexample set** $\Gamma, \neg\varphi$ (Definition 1.1.3) is true. Section 4.6 searched for a row; here is the search for a structure, the named move **search for a counter-model**.

1. Write the counterexample set, and push each $\neg$ inward (Section 6.2), so that every demand is visible as a $\forall$ or an $\exists$.
2. Meet the **existential demands first**: each $\exists x\, \psi$ asks for an element; try one already present, and add a new one only if every present one clashes.
3. Then impose each **universal constraint** on every element so far, deciding the atomic facts it forces.
4. **Add only forced elements.** If a constraint creates a new existential demand, return to step 2.
5. If every attempt at size n clashes, try size $n + 1$; if each new element forces another, look for a familiar infinite structure ($\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$) with that pattern.
6. **Check**: premises first (the warning above), then the conclusion, clause by clause.



The search for a counter-model. The procedure need not stop: Exercise C1 below gives a claim whose counter-models are all infinite, and Exercise C2 a sentence whose models are all infinite.

Example. A Counter-model, Found by the Procedure

Show that

$$\forall x\,(P(x) \to \exists y\,R(x,y)),\ \exists x\,P(x) \not\models \exists x\,\exists y\,(R(x,y) \wedge P(y)).$$

Solution. Step 1. The counterexample set, with the negated conclusion pushed inward:

$$\forall x\,(P(x) \to \exists y\,R(x,y)), \quad \exists x\,P(x), \quad \forall x\,\forall y\,(R(x,y) \to \neg P(y)).$$

Step 2. The demand $\exists x\,P(x)$ asks for an element: call it a, with $P(a)$.

Step 3. The first member, at a, now demands some y with $R(a,y)$.

Step 2 again. Try $y = a$, already present: then the third member, at $x = a$, $y = a$, forces $\neg P(a)$, a clash. So add b, with $R(a,b)$, and, by the third member, $\neg P(b)$.

Step 3 again. The constraints at b: the first member at b demands nothing, since $P(b)$ is false; the third member holds for every pair, since the only related pair is (a,b) and b lacks P. No new demand: stop adding.

Step 6. Take $D = \{a,b\}$, $P^{\mathfrak{M}} = \{a\}$, $R^{\mathfrak{M}} = \{(a,b)\}$. The first premise holds at a, with the witness b, and at b, whose hypothesis is false. The second holds at a. The conclusion fails: its only candidate pair (a,b) has $b \notin P^{\mathfrak{M}}$.

In numbers: on $\mathbb{N}$, read $P(x)$ as "x is even" and $R(x,y)$ as "$y = 2x+1$". Every even number x has the number $2x+1$ related to it, some number is even, and every number $2x+1$ is odd, so no pair (x,y) has $R(x,y)$ with y even.

Quick check. In the example, why was $y = a$ tried before adding b?

Solution. Step 2: a demand is met by an element already present when possible, since every new element brings new constraints. Here a clashed, so b was forced.

7.2.4 Small domains, by expansion

Over a domain with few elements, say $D = \{a, b\}$, the clauses can be unwound to the end. By (S6), $\mathfrak{M}, \ell \models \forall x\, \varphi$ exactly when φ is satisfied under $\ell[x \mapsto a]$ **and** under $\ell[x \mapsto b]$; by (S7), $\mathfrak{M}, \ell \models \exists x\, \varphi$ exactly when φ is satisfied under $\ell[x \mapsto a]$ **or** under $\ell[x \mapsto b]$; (S4) and (S5) treat the connectives as before; and at the bottom each atomic formula is decided by (S1) as a fact about elements, such as "the pair (a, b) is in $R^{\mathfrak{M}}$".

The **expansion** is the record of that check: $\forall x$ is written as a long $\wedge$ over the elements and $\exists x$ as a long $\vee$, the long $\wedge$ and $\vee$ of Section 5.1, and $R(a, b)$ is written for the fact $(a, b) \in R^{\mathfrak{M}}$. Here a and b are the elements themselves, not new constants, and no formula $\varphi[a/x]$ is formed: the expansion is notation for the finite table of (S6) and (S7) cases, so it is true exactly when the sentence is, by (S6), (S7) and (S1) alone. After the expansion only facts like $R(a, b)$ remain, each true or false, and the search for a counter-model of that size is a search for a row (Section 4.6).

Example. Size One Fails, Size Two Works

Find the least size of a counter-model to

$$\forall x\, \exists y\, R(x, y),\ \forall x\, \forall y\, \big(R(x, y) \to R(y, x)\big) \models \exists x\, R(x, x).$$

Solution. Size 1, $D = \{a\}$. The first premise expands to $R(a, a)$, which is also the conclusion's expansion. Whenever the first premise is true, so is the conclusion: no counter-model.

Size 2, $D = \{a, b\}$. The conclusion is false exactly when $\neg R(a, a) \wedge \neg R(b, b)$. The first premise expands to

$$\big(R(a, a) \vee R(a, b)\big) \wedge \big(R(b, a) \vee R(b, b)\big),$$

which then forces $R(a, b)$ and $R(b, a)$. The second premise expands to four implications, $R(a, a) \to R(a, a)$, $R(a, b) \to R(b, a)$, $R(b, a) \to R(a, b)$, $R(b, b) \to R(b, b)$, all true with $R^{\mathfrak{M}} = \{(a, b), (b, a)\}$.

So the least size is 2: two elements, each related to the other and neither to itself. In numbers: on $\{0, 1\}$, read $R(x, y)$ as $x \neq y$.

Remark. If a claim uses only one-place predicate symbols, k of them, and no constants, function symbols or $=$, and it has a counter-model at all, then it has one with at most 2^k elements: merge any two elements that have the same predicates. This book does not prove it. Once a two-place predicate symbol is allowed, no such bound exists: by Exercise C2 below, the claim $\models \neg\tau$, in which only the two-place symbol $<$ occurs, has counter-models (the models of τ, such as $(\mathbb{Z}, <)$), and every one of them is infinite. A function symbol with

$=$ does the same: the claim $\models \sigma$ of Exercise C1 has no finite counter-model at all.

Quick check. Expand $\exists x\, \forall y\, R(x, y)$ over $\{a, b\}$.

Solution. $(R(a,a) \wedge R(a,b)) \vee (R(b,a) \wedge R(b,b))$: the outer $\exists x$ becomes the $\vee$, and the inner $\forall y$ the $\wedge$ inside each disjunct.

7.2.5 Each side condition, broken

Chapter 5 introduced each side condition with a step that breaks it and a reading in which that step leads from truth to falsehood; here they are together, each reading now a counter-model.

Side condition	WRONG (breaks it)	RIGHT (keeps it)	Counter-model to the WRONG sequent
$\forall$E: t free for x (5.3)	from $\forall x\, \exists y\, (y \neq x)$, $\exists y\, (y \neq y)$	from the renamed premise $\forall x\, \exists z\, (z \neq x)$, $\exists z\, (z \neq y)$	$D = \{0, 1\}$: the premise holds; $\exists y\, (y \neq y)$ never does
$\forall$I: flag in no premise (5.4)	$P(x_0) \vdash \forall x\, P(x)$	$P(x_0) \vdash \exists x\, P(x)$, by $\exists$I	$\mathbb{Z}$, P "even", $\ell(x_0) = 2$
$\forall$I: flag in no open assumption (5.4)	$\vdash x_0 > 0 \to \forall x\, (x > 0)$	$\vdash \forall x\, (x > 0 \to x + 1 > 0)$, the boxes nested the other way	$\mathbb{Z}$, $\ell(x_0) = 1$
$\forall$I: flag not in the conclusion (5.4; 5.5, Exercise (C1: The conclusion clause of for-all-introduction, used later))	from $\forall x\, (x \leq x)$, a box flagged y_0 ending in $y_0 \leq y_0$ gives $\forall x\, (x \leq y_0)$, then $\exists y\, \forall x\, (x \leq y)$ by $\exists$I	from the same box, $\forall y\, (y \leq y)$	$\mathbb{N}$ with $\leq$: the premise holds, and no number is at least every number ($y + 1$ exceeds y)
$\exists$I: t free for x (5.5)	from $\forall y\, R(y, y)$, $\exists x\, \forall y\, R(x, y)$	from the same line, $\exists x\, R(x, x)$ ($\forall$E at x_0, then $\exists$I)	$\mathbb{Z}$, R as $\leq$: no least integer
$\exists$E: x_0 not in χ (5.6)	$\exists x\, (x > 5) \vdash \forall x\, (x > 5)$, through an $\exists$E line that keeps x_0	$\exists x\, (x > 5) \vdash \exists x\, (x > 6)$	$\mathbb{N}$: the premise holds at 6, the conclusion fails at 0
$\exists$E: x_0 in no premise or open assumption (5.6)	$\exists x\, P(x), \neg P(x_0) \vdash \bot$	$\exists x\, P(x), \forall x\, \neg P(x) \vdash \bot$	$D = \{1, 2\}$, $P^{\mathfrak{M}} = \{2\}$, $\ell(x_0) = 1$
$\exists$E: x_0 not in $\exists x\, \varphi$ (5.6)	$\forall y\, \exists x\, R(x, y) \vdash \exists z\, R(z, z)$	$\forall y\, \exists x\, R(x, y) \vdash \exists z\, \exists x\, R(x, z)$	three dots, $a \to b \to c \to a$
$=$E: terms free for x (5.7)	$a = y, \exists y\, (y < a) \vdash \exists y\, (y < y)$	$a = y, \exists z\, (z < a) \vdash \exists z\, (z < y)$, after renaming the premise	$\mathbb{Z}$, a and y both naming 0

Each counter-model makes the WRONG sequent's premises true and its conclusion false, so,

by the statement in the next subsection, no correct derivation reaches that conclusion: the side condition is not a fussy extra, it is what keeps the rule from leading from truth to falsehood. Each RIGHT step is legal, and its sequent has no counter-model, with one proviso. The RIGHT derivations of the rows "flag in no open assumption" and "x_0 not in χ", taken from Sections 5.4 and 5.6, use an "arith" line, which records an arithmetic fact and is not a rule. Those two sequents hold in every structure in which that fact ($\forall x\,(x > 0 \to x + 1 > 0)$, or $\forall x\,(x > 5 \to x + 1 > 6)$) is true, such as $\mathbb{N}$ or $\mathbb{Z}$ with the usual meanings, and not in every structure: read $+$ as subtraction on $\mathbb{Z}$, and $\forall x\,(x > 0 \to x + 1 > 0)$ fails at 1. Section 7.3 (optional) shows the same thing from the other side: each side condition is used in exactly one case of the soundness proof.

7.2.6 From counter-models to non-derivability

A counter-model shows $\Gamma \not\models \varphi$. The stalled searches asked something else: whether $\Gamma \vdash \varphi$. The bridge is the same as in Chapter 4, where Theorem 4.7.2 turned a refuting row into a proof that no derivation exists (Corollary 4.7.3). For predicate logic it reads:

Soundness for predicate logic. If $\Gamma \vdash \varphi$ with the rules of Chapters 1 and 2 and the quantifier and equality rules of Sections 5.3 to 5.7, then $\Gamma \models \varphi$ in the sense of Definition 7.1.4. Consequently, a counter-model to $\Gamma \models \varphi$ shows $\Gamma \nvdash \varphi$.

This statement is **not proved in this section**. Its proof extends Chapter 4's to the new rules, and it is the subject of Section 7.3, which is optional. The proof is where each side condition of the quantifier rules is used: the fresh-name conditions of $\forall$I and $\exists$E, and "free for" in $\forall$E, $\exists$I, and $=$E. A reader who skips Section 7.3 may take the statement on trust, as the rest of the book does; a reader who reads it will see why the side conditions are exactly as strict as they are.

With the statement, the examples above and the three pictures at the start of this section become non-derivability results. $\forall y\,\exists x\,R(x, y) \nvdash \exists x\,\forall y\,R(x, y)$: the converse search of Proposition 5.6.2 had to stall, and the two lines where it stalled, the blocked $\forall$I and the blocked close of the $\exists$E box, are not a lack of cleverness. Likewise the converse of Proposition 5.6.1, the converse of the $\forall/\vee$ law, $\exists x\,P(x) \vdash P(c)$, and $P(x_0) \vdash \forall x\,P(x)$ have no derivations. And the same reasoning is the semantic side of Section 6.4: there, Example (The square function on the real line) showed that x^2 on $\mathbb{R}$ is continuous and not uniformly continuous, so the arithmetic structure on $\mathbb{R}$, with f read as squaring, is a counter-model to "continuous entails uniformly continuous".

Quick check. Which row of the side-condition table does picture (C) belong to, if $\exists x\,P(x) \vdash P(c)$ is "derived" by an $\exists$E box with witness c?

Solution. None exactly: c is a constant, not a variable, so it cannot be the name of an $\exists$E box at all. With a variable x_0 in its place, the box would end in $P(x_0)$, and the derivation would break the row "x_0 not in χ". Either way, picture (C) shows that no derivation exists.

Warning. A counter-model proves **non**-derivability; it never proves derivability, and a structure in which a sentence is true says nothing about whether the sentence is derivable. $\forall x\,\exists y\,(x < y)$ is true in $(\mathbb{N}, <)$, yet it is not derivable: it is false in $(\{1, 2, 3\}, <)$. To show $\Gamma \vdash \varphi$, write a derivation.

What this certifies. Chapters 5 and 6 said, at every stall, that a counter-reading would settle the question and that Section 7.2 would explain why. Each of those readings is a structure, and the statement above turns each into a theorem:

- Section 5.1: the swapped pair of Example (Swapping two quantifiers), read over $\mathbb{R}$: $\forall x\,\exists y\,(x + y = 0)$ is true and $\exists y\,\forall x\,(x + y = 0)$ false, so the first does not derive the second (5.1's other refutation is a row, which Corollary 4.7.3 already settled in Chapter 4).
- Section 5.3: the three tiny worlds of the common errors (capture, a buried quantifier, some occurrences only), named in the paragraph on counter-readings, and the breaks of Exercise (B5: Derive or break).
- Section 5.4: $P(x_0) \vdash \forall x\,P(x)$ and the open-assumption error; the stalled converse of Exercise (B2: Contraposition under a quantifier, used later), $\forall x\,\neg Q(x) \to \forall x\,\neg P(x) \nvdash \forall x\,(P(x) \to Q(x))$, by the reading over $\mathbb{N}$ with P "even" and Q "odd" (the premise holds because its hypothesis $\forall x\,\neg Q(x)$ fails at 1; the conclusion fails at 0, even and not odd; Exercise B3 asks for a counter-model with exactly two elements); and the breaks (b), (d), (f) of Exercise (B5: Derive or break: properties of a relation, used later) (Section 7.4 makes (d) and (f) models).
- Section 5.5: $\exists x\,P(x) \vdash P(c)$ (picture (C)), the capture in $\exists$I (no least integer), $\forall x\,(x \leq x) \nvdash \exists y\,\forall x\,(x \leq y)$ (Exercise (C1: The conclusion clause of for-all-introduction, used later)), and the breaks of Exercise (B4: Derive or break).
- Section 5.6: the three counter-worlds of the $\exists$E clauses, the converse of Proposition 5.6.1 (picture (A)), the swap (picture (B), Example (The swap of $\exists$ and $\forall$)), and the breaks of Exercise (B4: Derive or break).
- Section 5.7: the capture in =E, the two-dot counter-reading to the converse of Example (Equal Sets Have the Same Elements) (the subsection "Extensionality" below), and the breaks of Exercise (B3: Derive or break).
- Section 5.8: the one-element domain with $0 \in 0$, in which any two sets without elements are equal and no set without elements exists (Exercise (B4: The empty set is unique) (b)).
- Section 6.1: the readings of the translation table, each separating the wrong formula from the right one. In rows (c)–(e) the right formula is true and the wrong one false, so the wrong one is not derivable from the right one; in rows (a) and (b) it is the other way round, so the right one is not derivable from the wrong one. Also the converse in Exercise (C3: Eventually, therefore infinitely often) (b).
- Section 6.2: the converses of (c) and (d) of Proposition 6.2.6 (6.2 read them over $\mathbb{Z}$; picture (A) and the first Quick check draw them), and the breaks of Exercise (B4: Derive or break).
- Section 6.3: every break of Exercise (B1: Derive or break, twelve claims) and Exercise (C2: Every M is P, every M is S, so some S is P?) (a), and the move "test the subgoal by a counter-reading".
- Section 6.4: continuity does not give uniform continuity (Example (The square function on the real line), as above), and the break in Exercise (C3: Exit check: a quantifier derivation) (b).

None of them was used as a reason before now; each is now a proof that no derivation exists.

7.2.7 Three classes of sentences

Fix the symbols of arithmetic: the constant 0, the one-place function symbol S, the two-place function symbols $+$ and $\cdot$, and the two-place predicate symbol $<$. A sentence in these symbols is

- an **instance of a tautology** if its propositional skeleton (Definition 6.3.1) is a tautology

(Definition 4.5.1);
- **valid** if it is true in every structure for these symbols ($\models \sigma$);
- **true in** $\mathbb{N}$ if it is true in the structure $(\mathbb{N}; 0, S, +, \cdot, <)$, with S the successor, $S(n) = n+1$.

Each class lies inside the next, and neither inclusion is an equality.

Example. Three Classes of Sentences

Place each sentence in the narrowest class it belongs to.

(a) $\forall x\,(x < S(x)) \vee \neg\forall x\,(x < S(x))$.
(b) $\forall x\,(x = x)$.
(c) $\forall x\,\neg(S(x) = 0)$.
(d) $\exists x\,(S(x) = 0)$.

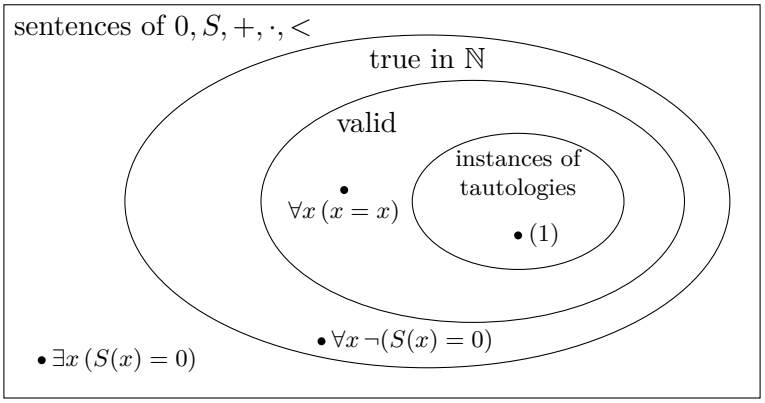
Solution. (a) An instance of a tautology: its skeleton is $p \vee \neg p$.

(b) Valid, by (S2) and (S6): every element equals itself. Not an instance of a tautology: its skeleton is one letter, p.

(c) True in $\mathbb{N}$: no successor is 0. Not valid: in the structure with $D = \{a\}$, $0^{\mathfrak{M}} = a$, $S^{\mathfrak{M}}(a) = a$ (and $+$, $\cdot$, $<$ read in any way), $S(0) = 0$ holds.

(d) False in $\mathbb{N}$, so in none of the three classes.

Why the inclusions hold: a valid sentence is true in $\mathbb{N}$ because $\mathbb{N}$ is one structure; an instance of a tautology whose different prime formulas have different letters (every sentence classified here) is valid because (S3)–(S5) compute a formula's truth from the truth of the prime formulas its skeleton's letters stand for, exactly as a truth table does (the first fold below gives the argument); that every instance is valid, including those where renamed prime formulas share a letter, is shown in the optional Section 7.3. The classes meet derivations as follows: an instance of a tautology is derivable (by TAUT, Proposition 6.3.1); a valid sentence need not be derivable by TAUT alone; and a sentence true in $\mathbb{N}$ but not valid is never derivable, by the statement of this section.



The three classes, each inside the next. The point (1) is $\forall x\,(x < S(x)) \vee \neg\forall x\,(x < S(x))$; the other three points are labeled with their sentences.

Quick check. Classify $\forall x\,(x < S(x)) \to (0 < S(0))$.

Solution. Valid: if $\forall x\,(x < S(x))$ holds in a structure, then by (S6) at $d = 0^{\mathfrak{M}}$ so does $0 < S(0)$, and (S5) makes the implication true; if it fails, (S5) makes the implication true anyway. Not an instance of a tautology: its skeleton is $p \to q$.

Second reading · *Why an instance of a tautology is valid*

Let σ be a sentence whose skeleton is a tautology and in which different prime formulas get different letters, and fix a structure $\mathfrak{M}$ and an assignment ℓ. Give each letter of the skeleton the truth value, under $\mathfrak{M}$ and ℓ, of the prime formula it stands for. This is one row of the skeleton's truth table. By induction on the skeleton, the skeleton's value in that row is T exactly when $\mathfrak{M}, \ell \models \sigma$: at a letter this is how the row was chosen, and at $\neg$, $\wedge$, $\vee$, $\to$, $\leftrightarrow$ the clauses (S3)–(S5) compute exactly as the clauses (V1)–(V6) of Definition 4.4.2 do. A tautology has value T in every row, so σ is satisfied in every structure under every assignment: it is valid.

When two different prime formulas share a letter (Definition 6.3.1 gives one letter to prime formulas that differ only by renaming bound variables), the same argument works, since the row above is well chosen only if the two formulas have the same truth value. They do, because renaming keeps meaning; the optional Section 7.3 proves this, and nothing in this section relies on it, since every sentence this section classifies gives different prime formulas different letters. The argument is about meaning and does not go through soundness.

Second reading · *A failed search proves nothing here, either*

A stalled derivation search proves nothing, as Chapters 5 and 6 said at every stall. A stalled counter-model search proves nothing either: the counter-models may all be infinite (Exercise C1), or simply larger than the sizes tried. Chapter 4's truth tables decided every propositional claim in finitely many rows (Proposition 4.6.1). For predicate logic, no mechanical procedure decides derivability: this is a theorem of Church and of Turing (1936), stated here without proof, and again in the optional Section 7.6. So the two searches are run side by side, and the reader's judgment chooses which to push.

7.2.8 Extensionality: an axiom, not a law of logic

Write $\in$ for a two-place predicate symbol. Section 5.7 derived that equal sets have the same members,

$$\vdash \forall x\,\forall y\,(x = y \to \forall z\,(z \in x \leftrightarrow z \in y))$$

(Example (Equal Sets Have the Same Elements)), and read its converse, **extensionality**,

$$\forall x\,\forall y\,(\forall z\,(z \in x \leftrightarrow z \in y) \to x = y),$$

in two dots with $\in$ holding nowhere. That reading is a structure: $D = \{a, b\}$, $\in^{\mathfrak{M}} = \emptyset$. Every z is a member of neither a nor b, so a and b have the same members, and $a \neq b$. By the statement of this section, extensionality is not derivable.

So extensionality is a fact about sets, not about logic: a theory of sets must take it as an axiom (Section 7.4 says what that means). The counter-model is not a strange set theory; it is a structure in which "$\in$" is read as a relation that holds nowhere.

Remark (Looking ahead). *Number Systems: From* $\mathbb{N}$ *to* $\mathbb{R}$, Chapter 1 takes extensionality as the first axiom of its set theory, (Z1), for exactly this reason: no derivation could supply it.

7.2.9 Exercises

7.2.9.1 A. Check your understanding

Exercise. A1

(a) Define a counter-model to $\Gamma \models \varphi$.

(b) State the soundness statement of this section, and say where it is proved.

(c) True or false, with a reason: if no structure with at most five elements is a counter-model to $\Gamma \models \varphi$, then $\Gamma \vdash \varphi$.

(d) True or false, with a reason: a structure in which the sentence σ is true shows $\vdash \sigma$.

(e) Which step of the search for a counter-model comes first, and which kind of demand is met first?

(f) Give, in one sentence, the reason that an instance of a tautology whose different prime formulas have different letters, such as (a) of Example (Three Classes of Sentences), is valid.

Solution. (a) A structure $\mathfrak{M}$ and an assignment ℓ with $\mathfrak{M}, \ell \models \Gamma$ and $\mathfrak{M}, \ell \not\models \varphi$.

(b) If $\Gamma \vdash \varphi$ with the rules of Chapters 1, 2 and 5, then $\Gamma \models \varphi$; so a counter-model shows $\Gamma \nvdash \varphi$. It is proved in the optional Section 7.3.

(c) False. The sentence σ of Exercise (C1: True in every finite model, used later) has no finite counter-model at all, and it is not derivable.

(d) False: a structure in which σ is true says nothing about derivability (the warning after the soundness statement). $\forall x\, \exists y\, (x < y)$ is true in $(\mathbb{N}, <)$ and not derivable.

(e) Writing the counterexample set with $\neg$ pushed inward comes first; existential demands are met before universal constraints are imposed.

(f) The clauses (S3)–(S5) compute truth in a structure from the truth of the prime formulas behind the skeleton's letters exactly as a row does, and a tautology is true in every row.

Exercise. A2: Drill

Each part takes a line or two. (a)–(c) give or judge a counter-model; (d) judges two infinite

ones; (e) and (f) complete a structure; (g) and (h) expand over $\{a, b\}$; (i)–(k) classify a sentence in the symbols 0, S, $+$, $\cdot$, $<$ as an instance of a tautology, valid but not an instance of a tautology, or true in $\mathbb{N}$ but not valid; (l) builds a counter-model.

(a) Give a counter-model to $\exists x\, \neg P(x) \models \neg \exists x\, P(x)$.

(b) Is $D = \{a\}$, $P^{\mathfrak{M}} = \{a\}$ a counter-model to $\forall x\, P(x) \models \exists x\, P(x)$?

(c) Give a counter-model to $\forall x\, (P(x) \to Q(x)) \models \forall x\, (Q(x) \to P(x))$.

(d) Is $\mathbb{Z}$, with $R(x, y)$ read as $x < y$, a counter-model to $\forall y\, \exists x\, R(x, y) \models \exists x\, \forall y\, R(x, y)$? And $\mathbb{N}$, with $R(x, y)$ read as $x > y$?

(e) $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$. Choose $Q^{\mathfrak{M}}$ so that the structure is a counter-model to $\forall x\, (P(x) \vee Q(x)) \models \forall x\, Q(x)$.

(f) $D = \{a, b\}$, and $(a, b) \in R^{\mathfrak{M}}$. Add pairs to $R^{\mathfrak{M}}$ so that the structure is a counter-model to $\forall x\, \exists y\, R(x, y) \models \exists y\, \forall x\, R(x, y)$.

(g) Expand $\forall x\, \exists y\, R(y, x)$ over $\{a, b\}$.

(h) Expand $\exists x\, (P(x) \wedge \forall y\, R(x, y))$ over $\{a, b\}$.

(i) $\forall x\, (x < x) \to \forall x\, (x < x)$.

(j) $\forall x\, (0 + x = x) \to (0 + 0 = 0)$.

(k) $\forall x\, \forall y\, (x + y = y + x)$.

(l) Give a counter-model to $\forall x\, (P(x) \to Q(x)),\ \exists x\, Q(x) \models \exists x\, P(x)$.

Solution.

(a) $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$: $\neg P$ holds at b, and $\exists x\, P(x)$ holds at a, so $\neg \exists x\, P(x)$ fails.

(b) No: both sentences are true there. And no structure is a counter-model, since domains are non-empty: if every element has P, some element does.

(c) $D = \{a\}$, $P^{\mathfrak{M}} = \emptyset$, $Q^{\mathfrak{M}} = \{a\}$: the premise holds since no element has P; the conclusion fails at a.

(d) Yes, both. In $\mathbb{Z}$ each y has $y - 1$ below it, and nothing is below everything (nothing is below itself). In $\mathbb{N}$ each y has $y + 1$ above it, and nothing is above everything (nothing is above itself).

(e) $Q^{\mathfrak{M}} = \{b\}$: it must contain b, which lacks P, and omit a, to break the conclusion.

(f) Add (b, a) and nothing else: b needs an arrow out, (b, b) would give b an arrow from every element, and (a, a) together with (b, a) would do the same for a.

(g) $\big(R(a, a) \vee R(b, a)\big) \wedge \big(R(a, b) \vee R(b, b)\big)$.

(h) $\big(P(a) \wedge R(a, a) \wedge R(a, b)\big) \vee \big(P(b) \wedge R(b, a) \wedge R(b, b)\big)$.

(i) An instance of a tautology: its skeleton is $p \to p$.

(j) Valid, by (S6) at $d = 0^{\mathfrak{M}}$ and (S5); not an instance of a tautology, since its skeleton is $p \to q$.

(k) True in $\mathbb{N}$, not valid: on $D = \{a, b\}$, read $x + y$ as "x" (and the other symbols in any way); then $a + b = a$ and $b + a = b$.

(l) $D = \{a\}$, $P^{\mathfrak{M}} = \emptyset$, $Q^{\mathfrak{M}} = \{a\}$: the first premise holds since no element has P, the second at a, and the conclusion fails.

7.2.9.2 B. Practice

Exercise. B1: Three counter-models, used later

Give a counter-model for each of the following, and check each premise and the conclusion against Definition 7.1.3.

(i) $\exists x\, P(x) \not\models \forall x\, P(x)$.

(ii) $\forall x\,(P(x) \to Q(x)) \not\models \exists x\, P(x) \to \forall x\, Q(x)$.

(iii) $\not\models \exists x\, P(x) \to P(x)$. (Here x is free in the last $P(x)$; the assignment is part of the answer.)

Solution. *(i)* Take $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$. The premise holds by (S7) with $d = a$. The conclusion fails by (S6) at $d = b$, since $b \notin P^{\mathfrak{M}}$.

(ii) Take $D = \{a, b\}$, $P^{\mathfrak{M}} = Q^{\mathfrak{M}} = \{a\}$. Premise: at $d = a$, $P(x)$ and $Q(x)$ are both satisfied, so $P(x) \to Q(x)$ is; at $d = b$, $P(x)$ is not satisfied, so $P(x) \to Q(x)$ is, by (S5). Hence the premise holds by (S6). Conclusion: $\exists x\, P(x)$ holds ($d = a$) and $\forall x\, Q(x)$ fails (at $d = b$), so by (S5) the implication fails. In words: whatever has P has Q, and something has P; that does not make everything have Q.

(iii) Take $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$, and an assignment with $\ell(x) = b$. The hypothesis $\exists x\, P(x)$ holds by (S7) with $d = a$; its x is bound, and the value $\ell(x)$ is overwritten. The conclusion $P(x)$ has x free, with value $\ell(x) = b \notin P^{\mathfrak{M}}$, so it fails. By (S5) the implication fails under ℓ. Under an assignment with $\ell(x) = a$ the implication holds; so the formula is satisfied under some assignments and not others, and a counter-model must name ℓ. Now take φ and ψ both to be $P(x)$. The formula $\forall x\,(P(x) \to P(x))$ is satisfied in every structure under every assignment, since $P(x) \to P(x)$ is satisfied at every d by (S5) and (S6). So this structure, with ℓ, is a counter-model to $\forall x\,(P(x) \to P(x)) \models \exists x\, P(x) \to P(x)$, and by the soundness statement of this section that sequent is not derivable. This is why the passage law $\forall x\,(\varphi \to \psi) \dashv\vdash \exists x\, \varphi \to \psi$ of Proposition 6.2.7 needs x not free in ψ.

Exercise. B2: Prove it or break it

For each claim $\Gamma \models \varphi$, either prove it by meaning (clause by clause or condensed) or give a counter-model, and state what follows about $\Gamma \vdash \varphi$. For (a) and (c), if the claim holds, also give a derivation.

(a) $\forall x\,(P(x) \to Q(x)),\ \forall x\, \neg Q(x) \models \forall x\, \neg P(x)$.

(b) $\forall x\,(P(x) \to Q(x)),\ \exists x\, \neg P(x) \models \exists x\, \neg Q(x)$.

(c) $\exists x \forall y\, R(x, y) \models \exists x\, R(x, x)$.

(d) $\forall x \exists y\, R(x, y) \models \forall x \exists y\, R(y, x)$.

(e) $\forall x\,(P(x) \to \exists y\, Q(y)) \models \exists x\, P(x) \to \exists y\, Q(y)$.

(f) $\exists x\, P(x) \to \forall y\, Q(y) \models \forall x\,(P(x) \to Q(x))$.

(g) $\forall x\,(P(x) \lor Q(x)),\ \exists x\, \neg P(x) \models \forall x\, Q(x)$.

(h) $\forall x\,(P(x) \lor Q(x)),\ \forall x\, \neg P(x) \models \forall x\, Q(x)$.

Solution. *(a) Holds.* If some d were in $P^{\mathfrak{M}}$, then $d \in Q^{\mathfrak{M}}$ by the first premise, against the second; so no d is in $P^{\mathfrak{M}}$. A derivation:

1		$\forall x\,(P(x) \to Q(x))$	
2		$\forall x\,\neg Q(x)$	
3	x_0	$P(x_0)$	asm, for $\neg$I
4		$P(x_0) \to Q(x_0)$	$\forall$E 1
5		$Q(x_0)$	$\to$E 4, 3
6		$\neg Q(x_0)$	$\forall$E 2
7		$\bot$	$\neg$E 5, 6
8		$\neg P(x_0)$	$\neg$I 3–7
9		$\forall x\,\neg P(x)$	$\forall$I 3–8

So $\Gamma \vdash \varphi$.

(b) Fails. $D = \{a\}$, $P^{\mathfrak{M}} = \emptyset$, $Q^{\mathfrak{M}} = \{a\}$: the first premise holds since its hypothesis is false at a, the second holds at a, and the conclusion fails since $a \in Q^{\mathfrak{M}}$. So $\Gamma \nvdash \varphi$, by the soundness statement.

(c) Holds. The premise's witness d has $(d, e) \in R^{\mathfrak{M}}$ for every e, in particular for $e = d$. A derivation:

1		$\exists x\,\forall y\,R(x, y)$	
2	x_0	$\forall y\,R(x_0, y)$	asm, for $\exists$E
3		$R(x_0, x_0)$	$\forall$E 2
4		$\exists x\,R(x, x)$	$\exists$I 3
5		$\exists x\,R(x, x)$	$\exists$E 1, 2–4

(d) Fails. $D = \{a, b\}$, $R^{\mathfrak{M}} = \{(a, b), (b, b)\}$: each element has an arrow out, and a has none in. So $\Gamma \nvdash \varphi$.

(e) Holds. If $\exists x\,P(x)$ is true, with witness d, the premise at d gives $\exists y\,Q(y)$; if it is false, the conclusion is true by (S5).

(f) Holds. If $d \in P^{\mathfrak{M}}$, then $\exists x\,P(x)$ holds, so $\forall y\,Q(y)$ holds by the premise, so $d \in Q^{\mathfrak{M}}$.

(g) Fails. $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$, $Q^{\mathfrak{M}} = \{b\}$: every element is in $P^{\mathfrak{M}}$ or $Q^{\mathfrak{M}}$, $b \notin P^{\mathfrak{M}}$, and $a \notin Q^{\mathfrak{M}}$. So $\Gamma \nvdash \varphi$.

(h) Holds. Each d is in $P^{\mathfrak{M}}$ or $Q^{\mathfrak{M}}$, and not in $P^{\mathfrak{M}}$; so it is in $Q^{\mathfrak{M}}$.

A proof by meaning shows $\Gamma \models \varphi$, not $\Gamma \vdash \varphi$. The derivations of (a) and (c) show $\vdash$; for (e), (f) and (h) the matching derivations exist but are not asked for (each is a short exercise with Chapter 5's rules). The pairs (a) and (b), and (g) and (h), are look-alikes; so are (c) and (d), with opposite prefixes: an $\exists\forall$ premise gives a loop, while a $\forall\exists$ premise does not even give every element an incoming arrow.

Marking scheme. (10 points)

- 8: 1 per part (a)–(h): the correct verdict with its justification, either a proof by meaning that covers every premise, or a counter-model with the premises checked true and the

conclusion checked false. A bare verdict scores 0.
- 1: the two derivations of (a) and (c), checked rule by rule (both needed).
- 1: stating, at a failing claim, that the counter-model gives $\Gamma \nvdash \varphi$ by the soundness statement.

Fatal errors (0 for that part): a "counter-model" that falsifies a premise; a proof by meaning offered as a derivation; a verdict given without a justification.

Exercise. B3: The converse from Section 5.4

Exercise (B2: Contraposition under a quantifier, used later) derived $\forall x\,(P(x) \to Q(x)) \vdash \forall x\,\neg Q(x) \to \forall x\,\neg P(x)$, and its search for the converse stalled; this section settled that converse above, in "What this certifies", with an infinite structure. Find a counter-model to

$$\forall x\,\neg Q(x) \to \forall x\,\neg P(x) \models \forall x\,(P(x) \to Q(x))$$

with exactly two elements, working backwards from the conclusion: say which element breaks the conclusion and how the premise is made true most cheaply.

Solution. Work backwards. The conclusion fails when some d has $d \in P^{\mathfrak{M}}$ and $d \notin Q^{\mathfrak{M}}$. The premise is cheapest to make true through a false hypothesis: make $\forall x\,\neg Q(x)$ false by putting some element in $Q^{\mathfrak{M}}$. Take $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$, $Q^{\mathfrak{M}} = \{b\}$.

Premise: $\forall x\,\neg Q(x)$ fails at $d = b$, since $b \in Q^{\mathfrak{M}}$; so the implication holds by (S5). Conclusion: at $d = a$, $a \in P^{\mathfrak{M}}$ and $a \notin Q^{\mathfrak{M}}$, so $P(x) \to Q(x)$ fails under $\ell[x \mapsto a]$, and (S6) fails. By the soundness statement of this section, the converse is not derivable, and the stalled search could not have succeeded.

Exercise. B4: Each side condition earns its place

Each derivation below breaks a side condition. For (a)–(c), name the line that breaks a side condition and the condition it breaks, give a counter-model to the sequent of its last line, and, where you can, repair the derivation so that it reaches a correct conclusion. For (d), say whether the conclusion is true (follows from the premise) and whether the derivation proves it.

(a)

1	x_0 | | $P(x_0)$		asm, for $\to$I
2	| | x_0 | $P(x_0)$		R 1
3	| | $\forall x\,P(x)$		$\forall$I 2–2
4	| $P(x_0) \to \forall x\,P(x)$		$\to$I 1–3
5	$\forall y\,(P(y) \to \forall x\,P(x))$		$\forall$I 1–4

(b)

1	$\forall y\,(y = y)$	
2	$\exists x\,\forall y\,(x = y)$	$\exists$I 1

(c)

1	$\exists x\,P(x)$	
2	$\forall x\,(P(x) \to Q(x))$	
3	x_0 | $P(x_0)$	asm, for $\exists$E
4	$P(x_0) \to Q(x_0)$	$\forall$E 2
5	$Q(x_0)$	$\to$E 4, 3
6	$Q(x_0)$	$\exists$E 1, 3–5

(d)

1	$P(x_0) \wedge \forall x\,P(x)$	
2	x_0 | $P(x_0)$	$\wedge$E 1
3	$\forall x\,P(x)$	$\forall$I 2–2

Solution. (a) Line 3: x_0 appears in the assumption on line 1, which is open at line 3. Line 5 is a legal $\forall$I: no premise, no open assumption, and x_0 is not in the conclusion. (This is the same-name nesting of Section 5.4's fold "The same error, as a false sentence", read there over $\mathbb{Z}$ with $x > 0$; here P is left open.) Counter-model to $\models \forall y\,(P(y) \to \forall x\,P(x))$: $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$. At $y = a$, $P(a)$ holds and $\forall x\,P(x)$ fails at b. No derivation reaches this conclusion: the counter-model shows it. A repair to a correct conclusion drops the inner box and uses $\exists$I in place of the illegal $\forall$I:

1	x_0 | | $P(x_0)$	asm, for $\to$I
2	| | $\exists x\,P(x)$	$\exists$I 1
3	| $P(x_0) \to \exists x\,P(x)$	$\to$I 1–2
4	$\forall y\,(P(y) \to \exists x\,P(x))$	$\forall$I 1–3

This is the RIGHT entry of the second row of the side-condition table, generalized.

(b) Line 2: in $\varphi = \forall y\,(x = y)$ the term y is not free for x, since it would be captured by $\forall y$. Counter-model: any domain with two elements; the premise holds, and no element equals both. Repairs: $\exists x\,\forall y\,(y = y)$ by $\exists$I from line 1, with x not in φ; or $z = z$ by $\forall$E at z, then $\exists x\,(x = z)$ by $\exists$I.

(c) Line 6: x_0 appears in $\chi = Q(x_0)$. The last line has x_0 free, so the counter-model needs an assignment: $D = \{a, b\}$, $P^{\mathfrak{M}} = Q^{\mathfrak{M}} = \{a\}$, $\ell(x_0) = b$. Both premises hold, and $Q(x_0)$ fails. Repair: forget the witness before the box closes.

1	$\exists x\, P(x)$		
2	$\forall x\,(P(x) \to Q(x))$		
3	x_0 | $P(x_0)$		asm, for $\exists$E
4		$P(x_0) \to Q(x_0)$	$\forall$E 2
5		$Q(x_0)$	$\to$E 4, 3
6		$\exists y\, Q(y)$	$\exists$I 5
7	$\exists y\, Q(y)$		$\exists$E 1, 3–6

(d) The conclusion is true: it follows from the premise, and is derived correctly in one line, $\forall x\, P(x)$ by $\wedge$E from line 1; so the sequent has no counter-model. The derivation does not prove it: line 3 breaks "the flag appears in no premise", since x_0 is in line 1. The counterexample is to the **step**, not to the conclusion: the same step takes $P(x_0)$ to $\forall x\, P(x)$, and $\mathbb{Z}$, with P read as "even" and $\ell(x_0) = 2$, makes $P(x_0)$ true and $\forall x\, P(x)$ false (the second row of the side-condition table).

Marking scheme. (10 points)

- 6: (a)–(c), 1 each for the line with its condition, and 1 each for a checked counter-model.
- 1: the repairs of (a)–(c), and the remark that (a)'s original conclusion is unreachable.
- 3: (d): the conclusion judged true, with the one-line derivation (1); the illegal line and its condition named (1); a counter-model to the step pattern rather than to the conclusion (1).

Fatal errors (0 for the step): in (c), a counter-model with the assignment omitted, although the last line has a free variable; in (d), a claimed counter-model to the conclusion.

7.2.9.3 C. Going deeper

Exercise. C1: True in every finite model, used later

Let f be a one-place function symbol, and let σ be the sentence

$$\forall x\, \forall y\, (f(x) = f(y) \to x = y) \to \forall y\, \exists x\, (f(x) = y),$$

"if f is injective, then f is surjective".

(a) Prove that σ is true in every structure whose domain is finite.

(b) Give a structure with an infinite domain in which σ is false.

(c) Deduce that σ is not derivable, although no finite structure is a counter-model.

Solution (Hint 1). For (a), suppose $f^{\mathfrak{M}}$ is injective on a finite D with n elements, and count the elements of its image.

Solution (Hint 2). For (b), look for an operation on $\mathbb{N}$ that never repeats a value but misses a value.

Solution. (a) Let $\mathfrak{M}$ have a finite domain D with $n \geq 1$ elements. If the hypothesis of σ is false in $\mathfrak{M}$, then σ is true by (S5). So suppose it is true: by (S6), (S5) and (S2), for all $d, e \in D$, $f^{\mathfrak{M}}(d) = f^{\mathfrak{M}}(e)$ implies $d = e$; that is, $f^{\mathfrak{M}}$ takes different values at different elements. Then the n elements of D have n different values, so the set of values $\{f^{\mathfrak{M}}(d) : d \in D\}$ is a subset of D with n elements. A subset of an n-element set with n elements is the whole set (a counting fact from school, proved in *Sets, Relations, and Cardinality*, Chapter 3). So every $e \in D$ is $f^{\mathfrak{M}}(d)$ for some d, which is the conclusion of σ by (S6), (S7) and (S2). Hence σ is true in $\mathfrak{M}$.

(b) Take $D = \mathbb{N}$ and $f^{\mathfrak{M}}(d) = d + 1$. The hypothesis holds: $d + 1 = e + 1$ gives $d = e$. The conclusion fails at $y = 0$: there is no $d \in \mathbb{N}$ with $d + 1 = 0$, since $d \geq 0$ gives $d + 1 \geq 1$. So σ is false in $\mathfrak{M}$.

(c) The structure of (b) is a counter-model to $\models \sigma$ (with Γ empty). By the soundness statement of this section, σ is not derivable. By (a), no finite structure could have shown this: a search for counter-models that lists finite structures of size 1, 2, 3, ... would never stop.

Marking scheme. (10 points)

- 1: (a) disposes of the case where the hypothesis of σ is false, by the clause for $\to$.
- 3: (a) reads the hypothesis as injectivity of $f^{\mathfrak{M}}$ and uses finiteness to get surjectivity, naming the counting fact.
- 3: (b) a structure with infinite domain, with the hypothesis checked and the conclusion refuted at a named element.
- 3: (c) concludes non-derivability through soundness and says why finite structures could not have done it.

Fatal errors (0 for the step): in (b), a "counter-model" in which $f^{\mathfrak{M}}$ is not injective; in (a), an argument that never uses finiteness.

Exercise. C2: A sentence with no finite model, used later

Let τ be the conjunction of the three sentences of the text saying that $<$ is a strict order with no largest element:

$$\forall x\, \neg(x < x) \;\wedge\; \forall x \forall y \forall z\, ((x < y \wedge y < z) \to x < z) \;\wedge\; \forall x\, \exists y\, (x < y).$$

Prove that every model of τ has an infinite domain. Where does your proof use that domains are non-empty?

Solution (Hint 1). Start from one element and keep climbing with the third conjunct.

Solution (Hint 2). Prove by induction on k that there are elements $d_1 < d_2 < \cdots < d_k$ of D with $d_i < d_j$ whenever $i < j$.

Solution. Let $\mathfrak{M}$ be a model of τ, with domain D, and write $d < e$ for $(d, e) \in <^{\mathfrak{M}}$. We prove, for every $k \geq 1$, the statement $A(k)$: *there are elements* $d_1, \ldots, d_k$ *of* D *with* $d_i < d_j$ *whenever* $1 \leq i < j \leq k$. We use Axiom 4.2.1, applied to the property $A(m+1)$ of $m \in \mathbb{N}$, so that the base case is $A(1)$.

Base. D is non-empty (Definition 7.1.1, (M1)), so there is some $d_1 \in D$, and $A(1)$ holds, with no pairs $i < j$ to check. This is the one place the proof uses non-emptiness. If an empty domain were allowed, with "for every $d \in \emptyset$" read as having nothing to check, all three conjuncts would hold in it (each begins with $\forall$), and it has no elements at all: the statement "every model is infinite" would fail.

Step. Suppose $d_1, \ldots, d_k$ are as in $A(k)$. By the third conjunct and (S6), (S7), there is $d_{k+1} \in D$ with $d_k < d_{k+1}$. For $i < k$, we have $d_i < d_k$ by $A(k)$, and then $d_i < d_{k+1}$ by the second conjunct (transitivity). So $A(k+1)$ holds.

Now the elements $d_1, \ldots, d_k$ are all different: if $d_i = d_j$ with $i < j$, then $d_i < d_j$ says $d_i < d_i$, which the first conjunct forbids. So for every $k \geq 1$, the domain contains k different elements. A finite set with n elements does not contain $n+1$ different elements, so D is not finite. Since $(\mathbb{N}, <)$ is a model of τ (as the text checked), τ has models, and all of them are infinite.

Marking scheme. (10 points)

- 2: names the non-empty domain as the start of the construction, and says what fails without it.
- 3: an induction producing, for every k, a chain of k elements, using the third conjunct for the step.
- 2: uses transitivity to keep the whole chain increasing, not only neighbors.
- 3: uses irreflexivity to show the elements are different, and concludes that D is infinite.

A second route: a proof that every non-empty finite subset of D has a $<$-maximal element (induction on its size, with irreflexivity in the base and transitivity in the step), applied to D, which is non-empty, and contradicted by the third conjunct, earns 2 for using D non-empty, 3 for the induction, 2 for transitivity in the step, and 3 for irreflexivity plus the contradiction with the third conjunct.

Fatal errors (0 for the step): building the chain without showing its elements are distinct; an "infinite chain" asserted without an induction or a recursion.

Exercise. C3: Exit check: formalize, negate, decide

Work over the domain $\mathbb{R}$, with $n \in \mathbb{Z}_{>0}$ as a bound as in Section 5.1, and $1/n$ as a term.

(a) Formalize, with bounded quantifiers: "for every $\varepsilon > 0$ there is $N \in \mathbb{Z}_{>0}$ such that $1/n < \varepsilon$ for every $n \in \mathbb{Z}_{>0}$ with $n \geq N$."

(b) Formalize the sentence obtained from (a) by moving "there is $N \in \mathbb{Z}_{>0}$" to the front.

(c) Negate each of (a) and (b) into positive form: no $\neg$ left in front of a quantifier or an implication, and $\neg(s < t)$ written as $s \geq t$.

(d) Decide which of (a) and (b) is true in $\mathbb{R}$. For the true sentence give a proof, saying what each chosen number depends on; for the false one, prove its negation, giving the witnesses it asks for.

You may use that for every real number there is a natural number larger than it.

Solution. A remark on the term $1/n$. In the structure $\mathbb{R}$, the symbol / must be a function on all of $\mathbb{R}^2$ (Definition 7.1.1, (M3)); give $x/0$ a value for every $x \in \mathbb{R}$, say $x/0 := 0$. The bound $n \in \mathbb{Z}_{>0}$ keeps $n \geq 1$, so those values are never used.

(a) $\forall \varepsilon > 0\ \exists N \in \mathbb{Z}_{>0}\ \forall n \in \mathbb{Z}_{>0}\ (n \geq N \to 1/n < \varepsilon)$.

(b) $\exists N \in \mathbb{Z}_{>0}\ \forall \varepsilon > 0\ \forall n \in \mathbb{Z}_{>0}\ (n \geq N \to 1/n < \varepsilon)$.

(c) Push $\neg$ inward one quantifier at a time, keeping each bound (Proposition 6.2.3), then negate the implication (Proposition 6.2.2):

$$\begin{aligned} \neg(a): &\quad \exists \varepsilon > 0\ \forall N \in \mathbb{Z}_{>0}\ \exists n \in \mathbb{Z}_{>0}\ (n \geq N \wedge 1/n \geq \varepsilon), \\ \neg(b): &\quad \forall N \in \mathbb{Z}_{>0}\ \exists \varepsilon > 0\ \exists n \in \mathbb{Z}_{>0}\ (n \geq N \wedge 1/n \geq \varepsilon). \end{aligned}$$

(d) **(a) is true.** Let $\varepsilon > 0$. Choose a natural number N with $N > 1/\varepsilon$, by the fact allowed in the problem; since $1/\varepsilon > 0$, $N \in \mathbb{Z}_{>0}$, and N depends on ε. Let $n \in \mathbb{Z}_{>0}$ with $n \geq N$. Then $n \geq N > 1/\varepsilon > 0$, and multiplying $n > 1/\varepsilon$ by the positive number ε/n gives $\varepsilon > 1/n$. Since n was arbitrary, and then ε, (a) holds.

(b) is false; its negation is true. Let $N \in \mathbb{Z}_{>0}$. Take $\varepsilon = 1/N$, which is positive, and $n = N$, which is in $\mathbb{Z}_{>0}$ with $n \geq N$. Then $1/n = 1/N \geq 1/N = \varepsilon$. Both witnesses depend on N, as the order $\forall N\ \exists \varepsilon\ \exists n$ allows. Since N was arbitrary, $\neg(b)$ holds, so (b) is false. In words: no single N works for every ε, because $\varepsilon = 1/N$ defeats that N.

The two sentences differ only in the order of $\forall \varepsilon$ and $\exists N$. In (a) the N is chosen after ε and may depend on it; in (b) it must be chosen first, which is the swap of Example (The swap of $\exists$ and $\forall$) in an analysis setting.

Marking scheme. (10 points)

- 2: both formalizations, with the bounds $\varepsilon > 0$, $N \in \mathbb{Z}_{>0}$, $n \in \mathbb{Z}_{>0}$ attached to the right quantifiers and $\to$ under the $\forall n$.
- 3: both negations in positive form, with every bound kept, the quantifier order of each preserved, and $\wedge$ under the $\exists n$.
- 2: verdict "true" for (a), with a proof in which N is chosen after ε and said to depend on it.
- 3: verdict "false" for (b), with ε and n given as functions of N and the three facts $\varepsilon > 0$, $n \geq N$, $1/n \geq \varepsilon$ checked.

Fatal errors (0 for the step): in the negation of (b), ε chosen before N (a single ε for all N); a bound negated instead of the quantifier ($\exists \varepsilon \leq 0$, or $\exists n < N$).

Exercise. C4: The least size of a counter-model

For each claim, find the least size of a counter-model, and prove by expansion that no smaller one exists.

(a) $\exists x\, P(x),\ \exists x\, Q(x),\ \exists x\, R(x) \models \exists x\, (P(x) \wedge Q(x)) \vee \exists x\, (Q(x) \wedge R(x)) \vee \exists x\, (P(x) \wedge R(x))$.

(b) $\forall x\, \exists y\, R(x, y),\ \forall x\, \neg R(x, x) \models \exists x\, \exists y\, (R(x, y) \wedge R(y, x))$.

(c) $\exists x\, P(x),\ \exists x\, \neg P(x),\ \exists x\, Q(x),\ \exists x\, \neg Q(x) \models \exists x\, (P(x) \wedge Q(x)) \vee \exists x\, (\neg P(x) \wedge \neg Q(x))$.

Solution (Hint). In each part, the premises demand witnesses. Count how many of them can share an element before the conclusion becomes true.

Solution. (a) Least size 3: $D = \{a, b, c\}$, $P^{\mathfrak{M}} = \{a\}$, $Q^{\mathfrak{M}} = \{b\}$, $R^{\mathfrak{M}} = \{c\}$. The premises hold with the witnesses a, b, c, and no element has two of the three predicates, so each disjunct of the conclusion fails. With at most two elements, the three witnesses for P, Q, R lie among two elements, so one element is a witness for two of them, and the matching disjunct holds. (Over $\{a, b\}$: the premises expand to $P(a) \vee P(b)$, $Q(a) \vee Q(b)$, $R(a) \vee R(b)$; of the three letters chosen, two fall on the same element.)

(b) Least size 3: $D = \{a, b, c\}$, $R^{\mathfrak{M}} = \{(a, b), (b, c), (c, a)\}$, a three-cycle: every element has an arrow out, none has a loop, and no two elements have arrows both ways. Size 1: the first premise expands to $R(a, a)$, against the second. Size 2: the second premise gives $\neg R(a, a) \wedge \neg R(b, b)$, so the first, $(R(a, a) \vee R(a, b)) \wedge (R(b, a) \vee R(b, b))$, forces $R(a, b)$ and $R(b, a)$, and the conclusion holds at $x = a$, $y = b$.

(c) Least size 2: $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$, $Q^{\mathfrak{M}} = \{b\}$: each element has exactly one of P, Q, so neither disjunct of the conclusion holds. Size 1: $\exists x\, P(x)$ and $\exists x\, \neg P(x)$ expand to $P(a)$ and $\neg P(a)$, which cannot both hold.

Marking scheme. (10 points)

- 4: (a): the counter-model checked (2); the counting argument for size 2 (2).
- 4: (b): the counter-model (1); size 1 (1); size 2 by expansion (2).
- 2: (c).

Fatal errors (0 for the step): a size claim without the proof that every smaller structure fails.

7.3 Soundness for Predicate Logic

This section is optional, and nothing outside it depends on it: Section 7.2 stated soundness for predicate logic without proof, and the rest of the book uses that statement. The results below build on one another, so they carry no optional marker of their own; the section as a whole is the unit a reader may skip. Here is the proof. It is Chapter 4's proof of Theorem 4.7.2 with six new cases, one for each quantifier and equality rule, and it is worth reading for one reason above all: each side condition of Sections 5.3 to 5.7 is used in the case of its own rule, and there one can see what would go wrong without it. The section proves a lemma about substitution and its corollary on renaming bound variables, then the theorem, then the corollary that turns every counter-model of Section 7.2 into a proof that no derivation exists.

7.3.1 What carries over, and what is new

Chapter 4 proved that whatever the rules of Chapters 1 and 2 derive is true in every row where the premises are true. The proof had four steps (Steps 1–4 of the proof of Theorem 4.7.2). It reduced to derivations in the primitive rules; it proved a claim for **every** line, not only the last, namely $\Gamma_k \models \varphi_k$, where Γ_k is the context of line k; it did so by strong induction on k, one case

per rule; and it read off the claim for the last line. We follow the same four steps. Step 2 is an induction, and the checklist of Section 4.2 applies to it as it applied to Chapter 4's: the claim stated for every line, the cases, and what each case may use. Three things change.

The formulas. Lines now carry first-order formulas (Definition 5.2.2), and $\models$ is the entailment of Definition 7.1.4: $\Gamma_k \models \varphi_k$ says that every structure $\mathfrak{M}$ and assignment ℓ with $\mathfrak{M}, \ell \models \Gamma_k$ satisfy φ_k. Assignments are not optional here. Lines inside a flagged box can mention the box's name, a **variable** such as x_0, free; such a line is then not a sentence, and its truth depends on the value of x_0. This is why Section 5.4 insisted that fresh names are always variables, never constants: the proof below changes their values.

The rules. The primitive rules are the thirteen of Chapter 4's list (Section 4.7) and six new ones: $\forall$E (Definition 5.3.1), $\forall$I (Definition 5.4.1), $\exists$I (Definition 5.5.1), $\exists$E (Definition 5.6.1), $=$I and $=$E (Definition 5.7.1). Two of the new rules cite a **box**, the flagged box of $\forall$I and the $\exists$E box, so the box rules are now seven: Chapter 4's five, and $\forall$I and $\exists$E. The other four new rules cite only lines and join the one-step rules.

↻ **Second reading** · *Derived rules and 'arith', written out*

Derived rules and "arith". As in Step 1 of Theorem 4.7.2, a line citing a derived rule abbreviates a pasted derivation in primitive rules. This covers Chapter 2's derived rules (Definition 2.3.1), whose derivations have no flagged boxes and apply to first-order formulas as they stand, and the bounded rules of Proposition 6.1.1, each of which Section 6.1 expanded into a displayed block of unbounded steps whose side conditions hold whenever the bounded ones do, and the one-line "(sym)" and "(trans)" of Section 5.7 (Proposition 5.7.1), each of which abbreviates its three-line derivation from $=$I and $=$E with the given terms pasted in. Those derivations have no boxes, and each φ they use is an equation, with no quantifier, so every term is free for its variable and the side conditions of Definition 5.7.1 hold. It also covers the shortcuts of Chapter 6. A law of Section 6.2 used as one line (QN, NIm, VI, the passage and distribution laws), and a derivable formula written as a line, abbreviate their printed derivations, pasted in with each flag renamed to a name new to the whole derivation, as "The laws as one line" (Section 6.2, after Proposition 6.2.1) argues. A TAUT line abbreviates the block of primitive steps that Proposition 6.3.1 provides, pasted in the same way (Section 6.3). In each case the side conditions of $\forall$I and $\exists$E still hold inside the pasted block. A line justified "arith" is not a rule at all: it records an arithmetic fact. Written out, it is a universally quantified arithmetic sentence taken as an extra **premise**, whose bound variables are distinct and appear nowhere else in the derivation, instantiated by $\forall$E at the terms on the line and applied by $\rightarrow$E, as Section 5.4 did explicitly when it turned the premises of a Chapter 1 derivation into universal statements. That choice of bound variables means that no fresh name of the derivation appears in the new premise, so the side conditions of $\forall$I and $\exists$E still hold, and that every term on the line is free for the variable it replaces at each $\forall$E step. So a derivation with "arith" lines is a derivation in primitive rules from its premises together with those arithmetic sentences, and soundness speaks about that derivation.

Renaming. Section 5.2 took, as a convention, that formulas differing only by renaming bound variables say the same thing, so a derivation may write a line, or apply a rule to a cited line, in a renamed form. This is not a rule either. Corollary 7.3.2 below shows that a formula and any renamed form of it are satisfied by exactly the same structures and assignments, so such a step changes nothing in the proof, which may take every line in the form its rule displays.

7.3.2 Contexts, once more

The context Γ_k of line k is still the list of the premises together with the assumptions of every box containing line k (Section 4.7, before Lemma 4.7.1). A flagged box of ∀I has **no** assumption, so it adds nothing to the context of the lines inside it. The ∃E box opens with its assumption $\varphi[x_0/x]$, which does join the context.

Section 5.4 extended Definition 1.4.2 to flagged boxes and kept its scope rule and its notion of availability word for word; two boxes are still disjoint or one inside the other. So Lemma 4.7.1 (a), whose proof uses only the scope rule, holds as it stands: if line i is available at line k, then $\Gamma_i \subseteq \Gamma_k$.

Part (b) needs one extension, since its statement names "the assumption of that box". It becomes: *if line k is justified by a box rule citing the box from line a to line b, then Γ_b is Γ_k together with the assumption of that box if it has one, and Γ_b is Γ_k if it is a flagged box without an assumption.* The proof of (b) in Section 4.7 is unchanged. It showed that the boxes containing line b are exactly the cited box and the boxes containing line k, and for that it used one fact about box rules: the last line of the cited box is a line of that box itself, not of a box inside it. The two new box rules demand the same: ∀I requires $\varphi[x_0/x]$ to be the last line of the flagged box, a line of that box itself (Definition 5.4.1), and ∃E requires χ to be the last line of the box itself (Definition 5.6.1). So the boxes containing line b are the cited box and those containing line k, and the context of line b adds to Γ_k only the cited box's assumption, if any. When a flagged box immediately contains another box and is cited from the first line of that inner box, as the citation convention of Section 5.4 says, the argument is the same: the cited box runs from that line to its last line, and its last line lies in no box inside it.

Here is the derivation of Proposition 5.6.2, with its contexts.

1	$\exists x\,\forall y\,R(x,y)$	
2	y_0 | x_0 | $\forall y\,R(x_0,y)$	asm, for ∃E
3	$R(x_0,y_0)$	∀E 2
4	$\exists x\,R(x,y_0)$	∃I 3
5	$\exists x\,R(x,y_0)$	∃E 1, 2–4
6	$\forall y\,\exists x\,R(x,y)$	∀I 2–5

Line	Boxes containing it	Context
1	none	$\exists x\,\forall y\,R(x,y)$
2, 3, 4	flagged y_0 (2–5); ∃E box x_0 (2–4)	$\exists x\,\forall y\,R(x,y)$, $\forall y\,R(x_0,y)$
5	flagged y_0 (2–5)	$\exists x\,\forall y\,R(x,y)$
6	none	$\exists x\,\forall y\,R(x,y)$

Line 6 cites the flagged box 2–5, which has no assumption, and indeed $\Gamma_5 = \Gamma_6$. Line 5 cites the ∃E box 2–4, and Γ_4 is Γ_5 with its assumption $\forall y\,R(x_0,y)$ added. Line 3 mentions x_0 and y_0 free, and line 4 mentions y_0 free, while its context mentions x_0 free, through the assumption $\forall y\,R(x_0,y)$: the claim $\Gamma_4 \models \exists x\,R(x,y_0)$ is about every structure **and every value** of x_0 and y_0 that makes the context true.

7.3.3 Substitution and satisfaction

Five of the six new rules write a substitution: ∀E and ∃I write $\varphi[t/x]$ for a term t, in their conclusion or premise; =E writes one in both; and ∀I and ∃E write $\varphi[x_0/x]$ for the fresh name. To handle them we need to know what $\varphi[t/x]$ **means** in a structure. The natural guess is that substituting the term t for x has the same effect as giving x the value of t. Section 5.2 showed that this can fail: in $\exists y\,(y > x)$, putting $y + 1$ for x blindly gives $\exists y\,(y > y + 1)$, which is false in $\mathbb{Z}$, while giving x the value of $y + 1$, whatever it is, leaves $\exists y\,(y > x)$ true. The y of the term was captured. The lemma says that capture is the only thing that can go wrong.

Lemma 7.3.1. Substitution Lemma

Let $\mathfrak{M}$ be a structure, ℓ an assignment in $\mathfrak{M}$, φ a formula, x a variable, and t a term that is free for x in φ (Definition 5.2.6). Write $a := t^{\mathfrak{M},\ell}$. Then

$$\mathfrak{M}, \ell \models \varphi[t/x] \quad \text{if and only if} \quad \mathfrak{M}, \ell[x \mapsto a] \models \varphi.$$

Idea. First terms: substituting t for x in a term s and then evaluating is the same as evaluating s with x set to the value of t. Then structural induction on φ, for all assignments at once, since the quantifier clauses change the assignment. Atomic formulas follow from the claim for terms, and connectives pass straight through. The only real case is $\forall y\,\varphi$ with x free in it. The substitution goes inside, and (S6) sets y to each d in turn. For the induction hypothesis to help, the value of t must not change when y does, and that is exactly what "free for" guarantees: y does not occur in t.

Proof. Fix $\mathfrak{M}$, x and t. For an assignment ℓ, write $a_\ell := t^{\mathfrak{M},\ell}$.

Terms. We show, for every term s and every assignment ℓ, that $(s[t/x])^{\mathfrak{M},\ell} = s^{\mathfrak{M},\ell[x\mapsto a_\ell]}$, by structural induction on s (Definition 5.2.1). If s is the variable x, then $s[t/x]$ is t, with value a_ℓ, and $s^{\mathfrak{M},\ell[x\mapsto a_\ell]} = \ell[x \mapsto a_\ell](x) = a_\ell$ (Definition 7.1.2). If s is another variable y, then $s[t/x]$ is y, and both sides are $\ell(y)$, since $\ell[x \mapsto a_\ell]$ changes only x. If s is a constant c, both sides are $c^{\mathfrak{M}}$. If s is $f(s_1, \ldots, s_n)$, then $s[t/x]$ is $f(s_1[t/x], \ldots, s_n[t/x])$ (Definition 5.2.5), and applying $f^{\mathfrak{M}}$ to the values, which agree by the hypothesis for $s_1, \ldots, s_n$, gives equal results.

Formulas. We prove, by structural induction on φ, the property $B(\varphi)$: *if t is free for x in φ, then for every assignment ℓ, $\mathfrak{M}, \ell \models \varphi[t/x]$ if and only if $\mathfrak{M}, \ell[x \mapsto a_\ell] \models \varphi$.* Two facts are used throughout. If t is free for x in $\neg\varphi$ or in $\varphi \circ \psi$, it is free for x in the parts, by the recursive description of "free for" after Definition 5.2.6 in Section 5.2. And by Definition 5.2.5, substitution passes through connectives: $(\neg\varphi)[t/x]$ is $\neg(\varphi[t/x])$, and $(\varphi \circ \psi)[t/x]$ is $\varphi[t/x] \circ \psi[t/x]$.

Atomic formulas. $R(s_1, \ldots, s_n)[t/x]$ is $R(s_1[t/x], \ldots, s_n[t/x])$. By the claim for terms, the tuple of values of the $s_i[t/x]$ under ℓ is the tuple of values of the s_i under $\ell[x \mapsto a_\ell]$, so (S1) of Definition 7.1.3 gives the same answer on both sides. Equations $s_1 = s_2$ are the same with (S2). For $\bot$, $\bot[t/x]$ is $\bot$, and neither side holds (S3).

Connectives. For $\varphi \wedge \psi$: by $B(\varphi)$ and $B(\psi)$, $\mathfrak{M}, \ell \models \varphi[t/x]$ if and only if $\mathfrak{M}, \ell[x \mapsto a_\ell] \models \varphi$, and likewise for ψ; (S5) combines the two answers in the same way on both sides. The cases $\neg$, $\vee$, $\to$, $\leftrightarrow$ are the same with (S4) and the other parts of (S5).

Universal quantifier, x not free in $\forall y\,\varphi$. This happens when y is the variable x, or when x is not free in φ. Then $(\forall y\,\varphi)[t/x]$ is $\forall y\,\varphi$ itself, since there is no free occurrence of x to replace (Section 5.2 proved that $\psi[t/x] = \psi$ whenever x is not free in ψ). The assignments ℓ and $\ell[x \mapsto a_\ell]$ differ at most at x, which is not free in $\forall y\,\varphi$; so by Lemma 7.1.1, $\mathfrak{M}, \ell \models \forall y\,\varphi$ if and only if $\mathfrak{M}, \ell[x \mapsto a_\ell] \models \forall y\,\varphi$, which is $B(\forall y\,\varphi)$.

Universal quantifier, x free in $\forall y\,\varphi$. Then y is not x, and x is free in φ. Every free occurrence of x in $\forall y\,\varphi$ lies in the scope of $\forall y$, so, as t is free for x in $\forall y\,\varphi$, **the variable y does not occur in t**, and t is free for x in φ (Definition 5.2.6, and its recursive description in Section 5.2). By Definition 5.2.5, $(\forall y\,\varphi)[t/x]$ is $\forall y\,(\varphi[t/x])$. Let $d \in D$. Since y does not occur in t, the assignments ℓ and $\ell[y \mapsto d]$ agree on every variable of t, so $a_{\ell[y\mapsto d]} = a_\ell$, by the claim for terms in the proof of Lemma 7.1.1. Since x and y are different variables, $\ell[y \mapsto d][x \mapsto a_\ell] = \ell[x \mapsto a_\ell][y \mapsto d]$. Hence, using $B(\varphi)$ for the assignment $\ell[y \mapsto d]$ in the second step,

$$\begin{aligned}
\mathfrak{M}, \ell \models \forall y\,(\varphi[t/x]) &\iff \text{for every } d \in D,\ \mathfrak{M}, \ell[y \mapsto d] \models \varphi[t/x] \\
&\iff \text{for every } d \in D,\ \mathfrak{M}, \ell[y \mapsto d][x \mapsto a_\ell] \models \varphi \\
&\iff \text{for every } d \in D,\ \mathfrak{M}, \ell[x \mapsto a_\ell][y \mapsto d] \models \varphi \\
&\iff \mathfrak{M}, \ell[x \mapsto a_\ell] \models \forall y\,\varphi,
\end{aligned}$$

where the first and last steps are (S6). This is $B(\forall y\,\varphi)$.

Existential quantifier. The same two subcases, with (S7) in place of (S6) and "for some d" in place of "for every d".

Every clause of Definition 5.2.2 is covered, so $B(\varphi)$ holds for every formula. This proves the lemma.

■

The hypothesis "free for" was used once, in the second quantifier case, to get $a_{\ell[y\mapsto d]} = a_\ell$. The capture example shows what happens without it. In the structure on $\mathbb{Z}$ with $>$, $+$, and 1 as usual, take $\varphi = \exists y\,(y > x)$ and $t = y + 1$, which is **not** free for x in φ. The blind substitution is $\exists y\,(y > y + 1)$, false under every assignment. But $\mathfrak{M}, \ell[x \mapsto a] \models \exists y\,(y > x)$ holds for every $a \in \mathbb{Z}$, with witness $a + 1$. The two sides of the lemma disagree, and the reason is visible in the displayed chain: when (S7) sets y to a new value, the value of $y + 1$ moves with it.

Quick check. Let $\mathfrak{M}$ be the structure on $\mathbb{Z}$ with $<$, $+$, and 1 as usual, $\varphi = \forall z\,(x < z \to y < z)$, $t = y+1$, and let $\ell(x) = 3$, $\ell(y) = 5$. Is t free for x in φ? Compute both sides of Lemma 7.3.1.

Solution. Yes: the only quantifier in φ is $\forall z$, and z does not occur in $y+1$. Here $a = t^{\mathfrak{M},\ell} = 6$. The right side asks whether $\mathfrak{M}, \ell[x \mapsto 6] \models \varphi$: every $z > 6$ is > 5, so yes. The left side is $\varphi[t/x] = \forall z\,(y + 1 < z \to y < z)$ under ℓ: every $z > 6$ is > 5, so yes. Both sides hold, as the lemma says. (Under ℓ itself, φ is false: $z = 4$ has $3 < 4$ and not $5 < 4$. The lemma compares $\varphi[t/x]$ with φ at the **new** value of x, not at the old one.)

The lemma also pays a debt of Section 5.2, which took renaming of bound variables inside a formula as a convention and left its justification in terms of meaning to this section.

Corollary 7.3.2. Renaming Keeps Meaning

Let φ be a formula and y, z variables, with z not appearing in φ (Section 5.2). Let $\mathfrak{M}$ be a structure for the symbols of φ, with domain D.

(a) For every assignment ℓ, $\mathfrak{M}, \ell \models \forall y\, \varphi$ if and only if $\mathfrak{M}, \ell \models \forall z\, \varphi[z/y]$, and $\mathfrak{M}, \ell \models \exists y\, \varphi$ if and only if $\mathfrak{M}, \ell \models \exists z\, \varphi[z/y]$.

(b) If ψ' arises from a formula ψ by such a renaming of one quantifier anywhere inside ψ (Section 5.2), then for every structure $\mathfrak{M}$ for the symbols of ψ and every assignment ℓ in it, $\mathfrak{M}, \ell \models \psi$ if and only if $\mathfrak{M}, \ell \models \psi'$.

Proof. *(a)* If z is y, then $\varphi[y/y]$ is φ and there is nothing to prove. Let z be a variable other than y. Since z does not appear in φ, no quantifier of φ is on z, so z is free for y in φ (Definition 5.2.6). Let $d \in D$. Under $\ell[z \mapsto d]$ the term z has value d, so Lemma 7.3.1 gives: $\mathfrak{M}, \ell[z \mapsto d] \models \varphi[z/y]$ if and only if $\mathfrak{M}, \ell[z \mapsto d][y \mapsto d] \models \varphi$. The assignments $\ell[z \mapsto d][y \mapsto d]$ and $\ell[y \mapsto d]$ differ at most at z, which does not appear in φ; by Lemma 7.1.1, the last statement holds if and only if $\mathfrak{M}, \ell[y \mapsto d] \models \varphi$. So $\mathfrak{M}, \ell[z \mapsto d] \models \varphi[z/y]$ if and only if $\mathfrak{M}, \ell[y \mapsto d] \models \varphi$, for every $d \in D$. Now (S6) gives the statement for $\forall$, and (S7) the statement for $\exists$.

(b) By structural induction on ψ, for all assignments at once, since the quantifier clauses change the assignment. If the renamed quantifier is the outermost symbol of ψ, this is (a). Otherwise ψ is not atomic, and the renaming takes place inside one immediate part of ψ, which it turns into a formula satisfied by the same assignments, by the induction hypothesis; the other parts are unchanged. If ψ is $\neg\alpha$ or $\alpha \circ \beta$, (S4) or (S5) then gives the same answer for ψ and ψ' under every ℓ. If ψ is $\forall w\, \alpha$ or $\exists w\, \alpha$, with the renaming inside α, then ψ' is $\forall w\, \alpha'$ or $\exists w\, \alpha'$, and α, α' are satisfied by the same assignments $\ell[w \mapsto d]$, so (S6) or (S7) gives the same answer for ψ and ψ'.

■

Several renamings one after another are covered by applying (b) once for each.

7.3.4 The soundness theorem

We can now state the theorem that Section 7.2 took on trust.

Theorem 7.3.3. Soundness for Predicate Logic

Let Γ be a list of formulas and ψ a formula. If $\Gamma \vdash \psi$ with the rules of Chapters 1 and 2 and the quantifier and equality rules of Sections 5.3 to 5.7, then $\Gamma \models \psi$ (Definition 7.1.4).

Idea. The steps are those of Chapter 4's proof (Steps 1–4 of Theorem 4.7.2).

① Reduce to derivations in the primitive rules.

② Prove $\Gamma_k \models \varphi_k$ for every line k, by strong induction on k (Theorem 4.2.2): fix a structure $\mathfrak{M}$ and an assignment ℓ satisfying Γ_k, and show that they satisfy φ_k.

③ The cases. Premises, assumptions, and the thirteen propositional rules go through word

for word, because (S3)–(S5) are Chapter 4's clauses. ∀E, ∃I, and =E are Lemma 7.3.1, used once or twice. =I is (S2). The two box rules are the heart of the proof, and they both use the same trick: **change the value of the fresh name**. In ∀I, to show $\forall x\,\varphi$ under ℓ, take any d, and set x_0 to d; the context does not notice, because x_0 is not in it, so the box's last line $\varphi[x_0/x]$ holds there, and moving d from x_0 to x gives φ under $\ell[x \mapsto d]$. In ∃E, set x_0 to the witness d that $\exists x\,\varphi$ provides; the box's assumption holds there, so its last line χ holds there, and since χ does not mention x_0, it holds under ℓ as well.

④ The last line lies in no box, so its context is the premises.

Before reading on: in the ∀I *case below, the assignment* $\ell[x_0 \mapsto d]$ *must still satisfy the context* Γ_k*, and at the end the conclusion must come back from* $\ell[x_0 \mapsto d][x \mapsto d]$ *to* $\ell[x \mapsto d]$*. Which clause of the side condition of* ∀I *does each step use?*

Proof. **Step 1 (reduction).** Suppose $\Gamma \vdash \psi$. By the subsection "What carries over, and what is new", there is a derivation of ψ in the primitive rules whose premises all belong to Γ. (A derivation with "arith" lines is of this kind only for the list of its premises together with the arithmetic sentences A those lines stand for, as explained there. The theorem applies to it with the list Γ, A in place of Γ, and gives $\Gamma, A \models \psi$: truth in every structure that satisfies Γ and A. The sentences A hold, for instance, when the arithmetic symbols have their usual meaning on $\mathbb{R}$, as in Section 7.1.) Fix one, with lines $1, \dots, n$, formulas φ_k, and contexts Γ_k. Let Σ be the finite set of constants, function symbols and predicate symbols that occur in some line of it. A line may contain symbols that occur neither in Γ nor in ψ, as c does in the line $P(c)$ between $\forall x\,P(x)$ and $\exists x\,P(x)$; so the induction is run over structures that interpret every symbol of Σ, and Step 4 returns to the structures of Definition 7.1.4.

Step 2 (the induction). For $k \in \mathbb{N}$, let $S(k)$ be: *if* $1 \le k \le n$*, then every structure* $\mathfrak{M}$ *that interprets every symbol of* Σ *and every assignment* ℓ *in* $\mathfrak{M}$ *with* $\mathfrak{M}, \ell \models \Gamma_k$ *satisfy* $\mathfrak{M}, \ell \models \varphi_k$. In Steps 2 and 3, "structure" means a structure that interprets every symbol of Σ, and possibly more. We prove $S(k)$ for all k by Theorem 4.2.2. For $k = 0$ and $k > n$, $S(k)$ holds vacuously. Let $1 \le k \le n$ and assume $S(j)$ for every $j < k$. Every line and box that line k cites is available at line k, so lies before it (Definition 1.4.2); hence the hypothesis holds for every cited line and for the last line of every cited box. Let $\mathfrak{M}$ be a structure, with domain D, and ℓ an assignment with $\mathfrak{M}, \ell \models \Gamma_k$. We show $\mathfrak{M}, \ell \models \varphi_k$. Two consequences of the hypothesis do the bookkeeping, as in the proof of Theorem 4.7.2. If line k cites line i, then $\Gamma_i \subseteq \Gamma_k$ by Lemma 4.7.1 (a), so $\mathfrak{M}, \ell \models \Gamma_i$, and $S(i)$ gives

$$\mathfrak{M}, \ell \models \varphi_i \quad \text{for every line } i \text{ cited by line } k. \tag{$*$}$$

If line k cites, by a box rule, the box from line a to line b, then by Lemma 4.7.1 (b), extended in the subsection "Contexts, once more", and $S(b)$:

every structure and assignment satisfying Γ_k and the box's assumption, if any, satisfy φ_b. (∗∗)

Step 3 (the cases).

Claim. If line k is a premise, an assumption, or justified by one of the thirteen primitive rules of Chapters 1 and 2, then $\mathfrak{M}, \ell \models \varphi_k$.

Proof of Claim. Claims 1–3 in the proof of Theorem 4.7.2 treat these cases. Each step there uses only (∗), (∗∗), the fact that a premise or an assumption belongs to its own context, and

the clauses (V1)–(V6) of Definition 4.4.2. Clauses (S3)–(S5) of Definition 7.1.3 are (V1)–(V6) with "v makes φ true" replaced by "$\mathfrak{M}, \ell \models \varphi$". Making that replacement throughout, the proofs of those claims prove this one, case by case.

■

Claim. If line k is justified by ∀E, ∃I, =I, or =E, then $\mathfrak{M}, \ell \models \varphi_k$.

Proof of Claim. ∀E. Line k is $\varphi[t/x]$, citing a line holding $\forall x\, \varphi$, with t free for x in φ. By $(*)$, $\mathfrak{M}, \ell \models \forall x\, \varphi$, so by (S6), $\mathfrak{M}, \ell[x \mapsto d] \models \varphi$ for every $d \in D$, in particular for $d = t^{\mathfrak{M},\ell}$. By Lemma 7.3.1, which applies because t is free for x in φ, $\mathfrak{M}, \ell \models \varphi[t/x]$.

∃I. Line k is $\exists x\, \varphi$, citing a line holding $\varphi[t/x]$, with t free for x in φ. By $(*)$, $\mathfrak{M}, \ell \models \varphi[t/x]$. By Lemma 7.3.1, $\mathfrak{M}, \ell[x \mapsto d] \models \varphi$ with $d = t^{\mathfrak{M},\ell} \in D$. So d is the element (S7) asks for, and $\mathfrak{M}, \ell \models \exists x\, \varphi$.

=I. Line k is $t = t$. The two sides have the same value $t^{\mathfrak{M},\ell}$, so (S2) holds.

=E. Line k is $\varphi[t_2/x]$, citing lines holding $t_1 = t_2$ and $\varphi[t_1/x]$, with t_1 and t_2 free for x in φ. By $(*)$ and (S2), $t_1^{\mathfrak{M},\ell}$ and $t_2^{\mathfrak{M},\ell}$ are the same element d; and $\mathfrak{M}, \ell \models \varphi[t_1/x]$. By Lemma 7.3.1 for t_1, $\mathfrak{M}, \ell[x \mapsto d] \models \varphi$. By Lemma 7.3.1 for t_2, read from right to left, $\mathfrak{M}, \ell \models \varphi[t_2/x]$.

■

Claim. If line k is justified by ∀I or ∃E, then $\mathfrak{M}, \ell \models \varphi_k$.

Proof of Claim. ∀I. Line k is $\forall x\, \varphi$, citing a flagged box with name x_0 and no assumption, whose last line b holds $\varphi[x_0/x]$. The side condition of Definition 5.4.1 says that at line k the name x_0 appears in no premise and in no assumption of a box still open at line k, and not in $\forall x\, \varphi$. The premises and the assumptions of the boxes open at line k are exactly the formulas of Γ_k, so

(i) x_0 appears in no formula of Γ_k; and

(ii) x_0 does not appear in $\forall x\, \varphi$, so x_0 is not the variable x, and x_0 does not appear in φ.

The flagged box has no assumption, so $(**)$ says: every structure and assignment satisfying Γ_k satisfy $\varphi[x_0/x]$.

Let $d \in D$; we show $\mathfrak{M}, \ell[x \mapsto d] \models \varphi$. Put $\ell' := \ell[x_0 \mapsto d]$. It differs from ℓ at most at x_0, which by (i) is free in no formula of Γ_k; so by Lemma 7.1.1, $\mathfrak{M}, \ell' \models \Gamma_k$. **This is where the fresh-name condition is used:** changing the value of x_0 cannot disturb anything we assumed. By $(**)$, $\mathfrak{M}, \ell' \models \varphi[x_0/x]$. By (ii), x_0 does not appear in φ, so no quantifier of φ is on x_0, and the term x_0 is free for x in φ (Definition 5.2.6). Its value under ℓ' is d, so Lemma 7.3.1 gives

$$\mathfrak{M}, \ell[x_0 \mapsto d][x \mapsto d] \models \varphi.$$

The assignments $\ell[x_0 \mapsto d][x \mapsto d]$ and $\ell[x \mapsto d]$ agree at x, where both give d, and at every variable other than x and x_0, where both give the value ℓ gives. They can differ only at x_0, which by (ii) does not appear in φ. By Lemma 7.1.1, $\mathfrak{M}, \ell[x \mapsto d] \models \varphi$. Since $d \in D$ was arbitrary, (S6) gives $\mathfrak{M}, \ell \models \forall x\, \varphi$.

∃E. Line k is χ, citing a line h holding $\exists x\, \varphi$ and a box with name x_0 that opens with the assumption $\varphi[x_0/x]$ and ends with χ. The side condition of Definition 5.6.1 says that x_0 appears neither in χ, nor in $\exists x\, \varphi$, nor in any premise, nor in any assumption of a box still open at line k. So

(i) x_0 appears in no formula of Γ_k;

(ii) x_0 is not x, and does not appear in φ;

(iii) x_0 does not appear in χ.

By $(**)$, every structure and assignment satisfying Γ_k and $\varphi[x_0/x]$ satisfy χ.

By $(*)$ for line h, $\mathfrak{M}, \ell \models \exists x\, \varphi$; by (S7), choose $d \in D$ with $\mathfrak{M}, \ell[x \mapsto d] \models \varphi$. Put $\ell' := \ell[x_0 \mapsto d]$. First, $\mathfrak{M}, \ell' \models \Gamma_k$, by (i) and Lemma 7.1.1, as in the $\forall$I case. Second, $\mathfrak{M}, \ell' \models \varphi[x_0/x]$. Indeed, $\ell'[x \mapsto d] = \ell[x_0 \mapsto d][x \mapsto d]$ and $\ell[x \mapsto d]$ differ at most at x_0, which by (ii) does not appear in φ, so $\mathfrak{M}, \ell'[x \mapsto d] \models \varphi$ by Lemma 7.1.1; the term x_0 is free for x in φ by (ii), and has value d under ℓ', so Lemma 7.3.1 gives $\mathfrak{M}, \ell' \models \varphi[x_0/x]$. **This is where "$x_0$ not in $\exists x\, \varphi$" is used.** By $(**)$, $\mathfrak{M}, \ell' \models \chi$. Finally, ℓ' and ℓ differ at most at x_0, which by (iii) does not appear in χ, so $\mathfrak{M}, \ell \models \chi$ by Lemma 7.1.1. **This is where "x_0 not in χ" is used.**

■

The three claims cover every way a line can be justified, so $S(k)$ holds, and by Theorem 4.2.2 $S(k)$ holds for every line k of the derivation.

Step 4 (the last line). The last line n lies in no box (Definition 1.4.2), so Γ_n consists of premises only, and $\varphi_n = \psi$. Let $\mathfrak{M}$ be a structure for the symbols occurring in Γ and ψ, with domain D, and ℓ an assignment with $\mathfrak{M}, \ell \models \Gamma$. Expand $\mathfrak{M}$ to a structure $\mathfrak{M}'$ with the same domain and the same meanings, which also interprets the symbols of Σ that $\mathfrak{M}$ does not, so that $\mathfrak{M}'$ interprets every symbol of Σ: fix some $e \in D$, which exists because D is non-empty (M1), and give each new constant the value e, each new function symbol, of arity m say, the function $D^m \to D$ with constant value e, and each new predicate symbol the empty relation. Then ℓ is an assignment in $\mathfrak{M}'$, and by the remark after Definition 7.1.4 (the clauses of Definition 7.1.3 consult only the symbols of the formula at hand), each formula of Γ, and ψ, is satisfied by $\mathfrak{M}', \ell$ exactly when it is satisfied by $\mathfrak{M}, \ell$. Every premise belongs to Γ, so $\mathfrak{M}', \ell \models \Gamma_n$; by $S(n)$, $\mathfrak{M}', \ell \models \psi$; and so $\mathfrak{M}, \ell \models \psi$. This proves $\Gamma \models \psi$.

■

The answer to the prompt: the first step uses the clauses "no premise" and "no open assumption", which together say that x_0 is not in Γ_k; the second uses "not in $\forall x\, \varphi$". The last line $\varphi[x_0/x]$ of the box, which may mention x_0 as often as it likes, plays no part in either. That is the precise sense in which "let x_0 be arbitrary" works: whatever value x_0 is given, what we know still holds, so what we proved about x_0 holds at that value.

The $\exists$E case uses the same move with the opposite purpose. There the value of x_0 is **not** arbitrary; it is set to the witness d. What makes that legitimate is that nothing else knows about x_0 (condition (i)), that the assumption $\varphi[x_0/x]$ then holds at the witness (condition (ii)), and that the result χ forgets it again (condition (iii)).

7.3.5 Counter-models prove non-derivability

The corollary is the one Section 7.2 used, now with a proof.

Corollary 7.3.4. Counter-models

Let Γ be a list of formulas and ψ a formula. If some structure $\mathfrak{M}$ and assignment ℓ satisfy

every formula of Γ and do not satisfy ψ, then $\Gamma \nvdash \psi$.

Proof. The pair $\mathfrak{M}, \ell$ shows that $\Gamma \models \psi$ fails (Definition 7.1.4). If $\Gamma \vdash \psi$ held, Theorem 7.3.3 would give $\Gamma \models \psi$. So $\Gamma \vdash \psi$ does not hold. ■

If the derivations in question may use "arith" lines, a counter-model must also satisfy the arithmetic sentences those lines stand for (Step 1 of the proof of Theorem 7.3.3), for instance by giving the arithmetic symbols their usual meaning.

Every counter-model of Section 7.2 is now a proof that no derivation exists.

- By Example (The swap of $\exists$ and $\forall$), $\forall y\, \exists x\, R(x, y) \nvdash \exists x\, \forall y\, R(x, y)$. The converse search of Proposition 5.6.2 stalled at two lines, a $\forall$I blocked by an open assumption and an $\exists$E blocked by its conclusion; now we know that no other route could have succeeded.
- The two-element structure with $P^{\mathfrak{M}} = \{a\}$, $Q^{\mathfrak{M}} = \{b\}$ gives $\exists x\, P(x) \wedge \exists x\, Q(x) \nvdash \exists x\, (P(x) \wedge Q(x))$ and $\forall x\, (P(x) \vee Q(x)) \nvdash \forall x\, P(x) \vee \forall x\, Q(x)$: the one-way laws of Proposition 6.2.6 are one-way.
- $\exists x\, P(x) \nvdash P(c)$, which Exercise (C2: A witness is not a name, used later) could only explain informally.
- $P(x_0) \nvdash \forall x\, P(x)$, by the structure on $\mathbb{Z}$ with $P^{\mathfrak{M}}$ the even integers and $\ell(x_0) = 2$: generalizing over a name in a premise is not merely forbidden, it would be unsound.
- The passage laws of Proposition 6.2.7 need x not free in ψ: with x free, $\nvdash \exists x\, P(x) \to P(x)$, by the counter-model of Exercise (B1: Three counter-models, used later) (iii), which names an assignment.

Remark. The converse of Theorem 7.3.3 is also true: if $\Gamma \models \psi$, then $\Gamma \vdash \psi$. This is the **completeness theorem** for predicate logic, proved by Gödel in 1929. Its proof needs to build a structure out of the formulas themselves, and it is beyond this book; nothing here depends on it.

Remark. One difference from Chapter 4 remains even with completeness: there is no truth table to check. Whether a sentence of predicate logic is derivable cannot, in general, be decided by any mechanical procedure (Church and Turing, 1936). A failed search and a missing counter-model can both go on for ever.

7.3.6 Every side condition is a case of the proof

The side conditions of Sections 5.3 to 5.7 were introduced one at a time, each with a counterexample. The proof above shows that they are all the same kind of thing: each is used in the case of its own rule, and dropping it breaks a step of that case. Each counterexample of Sections 5.3 to 5.7 can be read as a failed case:

- generalizing over a name in a premise, as in "$P(x_0) \vdash \forall x\, P(x)$" (Section 5.4): the $\forall$I case fails at "$\mathfrak{M}, \ell[x_0 \mapsto d] \models \Gamma_k$", since changing x_0 changes the premise $P(x_0)$;
- generalizing over a name in an open assumption, as in "from $x_0 > 0$, $\forall x\, (x > 0)$" (Section

5.4): the same step, with the assumption $x_0 > 0$ in Γ_k;
- a name that leaks into the conclusion of ∀I, as in Exercise (C1: The conclusion clause of for-all-introduction, used later): the last step of the ∀I case, since x_0 now occurs in φ;
- treating the witness as arbitrary, by ∀I over its name inside its ∃E box (Section 5.6): the ∀I case again, since the box's assumption is in Γ_k;
- letting the witness escape its box, or giving a second witness a name already in an open assumption (Section 5.6): the ∃E case, at condition (iii) or (i);
- a capture in ∀E (Section 5.3) or ∃I (Section 5.5): the ∀E or ∃I case, where Lemma 7.3.1 does not apply; Exercise C2 and the solution to Exercise B1 give false conclusions;
- a capture in =E, as in $a = y,\ \exists y\,(y < a) \vdash \exists y\,(y < y)$ (Section 5.7): the =E case, where Lemma 7.3.1 does not apply to t_2.

Warning. A side condition is not mere bookkeeping: a derivation that breaks one is not covered by Theorem 7.3.3, and it can reach a false line, as the counter-model to $P(x_0) \vdash \forall x\, P(x)$ in the previous subsection shows for the first item of the list.

Quick check. Exercise (C2: Find the error, used later) gives a derivation of $\exists x\,(P(x) \wedge Q(x)) \vdash \exists x\, P(x)$ whose ∃E box ends in $P(x_0)$; the line after the box is $P(x_0)$ by ∃E, and the last line is $\exists x\, P(x)$ by ∃I. Which case of the proof of Theorem 7.3.3 fails, at which step, and for which line? Give a structure and an assignment showing that the claim $\Gamma_k \models \varphi_k$ is false for that line.

Solution. The ∃E case fails, at its last step: $\chi = P(x_0)$ contains x_0, so condition (iii) is false, and "$\mathfrak{M}, \ell' \models \chi$" cannot be carried back from $\ell' = \ell[x_0 \mapsto d]$ to ℓ. The line concerned is the illegal ∃E line $P(x_0)$, whose context is the premise alone. Take $D = \{a, b\}$, $P^{\mathfrak{M}} = Q^{\mathfrak{M}} = \{a\}$ and $\ell(x_0) = b$: the premise $\exists x\,(P(x) \wedge Q(x))$ holds (witness a), and $P(x_0)$ fails, since $b \notin P^{\mathfrak{M}}$. The final conclusion $\exists x\, P(x)$ is still true in this structure, as it is in every structure satisfying the premise; the fault is in the line before it.

7.3.7 Exercises

7.3.7.1 A. Check your understanding

Exercise. A1

These parts are about the proof of Theorem 7.3.3.

(a) In the ∀I case, name the two places where the side condition of ∀I is used, and say which clause each uses.

(b) The derivation below breaks the side condition of ∃E: its conclusion $P(x_0)$ contains the fresh name. Which step of the ∃E case fails? Give a structure and an assignment in which the premise holds and line 3 does not.

(c) Give the idea of the proof in two sentences.

(d) Why does the claim $\Gamma_k \models \varphi_k$ quantify over assignments, and not only over structures? Point to a line of the derivation of Proposition 5.6.2 printed in this section.

1	$\exists x\, P(x)$	
2	$x_0 \mid P(x_0)$	asm, for $\exists$E
3	$P(x_0)$	$\exists$E 1, 2–2

Solution. *(a)* First, to show that $\ell[x_0 \mapsto d]$ still satisfies Γ_k: this uses "x_0 appears in no premise and in no assumption of a box open at line k", which together say that x_0 is in no formula of Γ_k, so Lemma 7.1.1 applies. Second, to pass from $\ell[x_0 \mapsto d][x \mapsto d]$ to $\ell[x \mapsto d]$ (and, before that, to know that x_0 is free for x in φ): this uses "x_0 does not appear in $\forall x\, \varphi$".

(b) The last step: from $\mathfrak{M}, \ell[x_0 \mapsto d] \models \chi$ to $\mathfrak{M}, \ell \models \chi$, which needs x_0 not in χ. Take $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$, $\ell(x_0) = b$. The premise $\exists x\, P(x)$ holds, with witness a; line 3, $P(x_0)$, fails, since $\ell(x_0) = b \notin P^{\mathfrak{M}}$. Inside the box, under $\ell[x_0 \mapsto a]$, line 2 holds; the proof would carry it back to ℓ, and that is exactly what fails.

(c) Every line of a derivation is satisfied by every structure and assignment that satisfy its context, by strong induction on the line, with one case per rule. The quantifier box rules are handled by changing the value of the fresh name, which the side conditions guarantee the context and the conclusion cannot notice.

(d) Lines inside boxes contain free variables, the fresh names, and their truth depends on the values of those variables. Line 3, $R(x_0, y_0)$, is not a sentence: in the structure on $\mathbb{Z}$ with R read as $<$, it is satisfied when $\ell(x_0) = 1$, $\ell(y_0) = 2$, and not when $\ell(x_0) = 2$, $\ell(y_0) = 1$. The $\forall$I and $\exists$E cases work precisely by changing such values.

7.3.7.2 B. Practice

Exercise. B1: The $\exists$I case

Write out the $\exists$I case of the proof of Theorem 7.3.3 in full: state what is known about the cited line, and show $\mathfrak{M}, \ell \models \exists x\, \varphi$, naming the reason for each step. Where is "t free for x in φ" used?

Solution. Line k is $\exists x\, \varphi$, by $\exists$I citing a line i holding $\varphi[t/x]$, where t is free for x in φ (Definition 5.5.1). Let $\mathfrak{M}$ be a structure interpreting every symbol of the derivation (as in Step 2), with domain D, and ℓ an assignment with $\mathfrak{M}, \ell \models \Gamma_k$. Line i is available at line k, so $\Gamma_i \subseteq \Gamma_k$ by Lemma 4.7.1 (a), and $\mathfrak{M}, \ell \models \Gamma_i$. Since $i < k$, the induction hypothesis $S(i)$ gives $\mathfrak{M}, \ell \models \varphi[t/x]$. Let $d := t^{\mathfrak{M}, \ell}$, an element of D. Because t is free for x in φ, Lemma 7.3.1 applies and gives $\mathfrak{M}, \ell[x \mapsto d] \models \varphi$. By (S7) of Definition 7.1.3, with the element d, $\mathfrak{M}, \ell \models \exists x\, \varphi$.

"Free for" is used exactly once, to apply Lemma 7.3.1. Without it the step fails: from $\forall y\, R(y, y)$, read as $(\forall y\, R(x, y))[y/x]$, an "$\exists$I" would give $\exists x \forall y\, R(x, y)$, and in the structure on $\mathbb{Z}$ with R read as $\leq$ the first is true and the second ("there is a least integer") false.

Exercise. B2: Three non-derivability results

Use the following counter-models from Section 7.2 to state three non-derivability results,

citing Corollary 7.3.4, and check in each case that the structure (and assignment) satisfy every premise and not the conclusion.

(a) The two-element structure of Example (The swap of $\exists$ and $\forall$).

(b) The counter-model of Exercise (B3: The converse from Section 5.4).

(c) The counter-model of Exercise (B1: Three counter-models, used later) (iii).

Solution. *(a)* $D = \{a, b\}$, $R^{\mathfrak{M}} = \{(a, a), (b, b)\}$. The premise $\forall y\, \exists x\, R(x, y)$ holds (for $y = a$ take $x = a$, for $y = b$ take $x = b$); the conclusion $\exists x\, \forall y\, R(x, y)$ fails ($x = a$ fails at $y = b$, $x = b$ at $y = a$). Both are sentences, so any assignment will do. By Corollary 7.3.4, $\forall y\, \exists x\, R(x, y) \nvdash \exists x\, \forall y\, R(x, y)$.

(b) $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$, $Q^{\mathfrak{M}} = \{b\}$. The premise $\forall x\, \neg Q(x) \to \forall x\, \neg P(x)$ holds, since its hypothesis fails at b; the conclusion $\forall x\, (P(x) \to Q(x))$ fails at a. By Corollary 7.3.4, $\forall x\, \neg Q(x) \to \forall x\, \neg P(x) \nvdash \forall x\, (P(x) \to Q(x))$.

(c) $D = \{a, b\}$, $P^{\mathfrak{M}} = \{a\}$, with an assignment $\ell(x) = b$. There are no premises, so the condition on premises holds trivially. The formula $\exists x\, P(x) \to P(x)$ fails under ℓ: its hypothesis holds (witness a) and its conclusion, with x free, has value $b \notin P^{\mathfrak{M}}$. By Corollary 7.3.4, with Γ empty, $\nvdash \exists x\, P(x) \to P(x)$. Here the assignment is essential: under $\ell(x) = a$ the formula holds, and the structure alone would not have been a counter-model.

7.3.7.3 C. Going deeper

Exercise. C1: The $\exists$E case

Write out the $\exists$E case of the proof of Theorem 7.3.3 in full, from memory: the data of line k, the side condition, what the induction hypothesis gives for the cited line and the cited box, and the argument, naming at each step which clause of the side condition it uses.

Solution (Hint 1). The tool is the assignment $\ell[x_0 \mapsto d]$, where d is a witness for $\exists x\, \varphi$ under ℓ. You will need Lemma 7.1.1 three times and Lemma 7.3.1 once.

Solution (Hint 2). Start: "Let $\mathfrak{M}$ be a structure interpreting every symbol of the derivation, with domain D, and ℓ an assignment with $\mathfrak{M}, \ell \models \Gamma_k$. By the induction hypothesis for line h and Lemma 4.7.1 (a), $\mathfrak{M}, \ell \models \exists x\, \varphi$, so choose $d \in D$ with $\mathfrak{M}, \ell[x \mapsto d] \models \varphi$."

Solution. *Data.* Line k is χ, by $\exists$E citing a line h holding $\exists x\, \varphi$ and a box with name x_0, from line a (the assumption $\varphi[x_0/x]$) to line b (the formula χ). Side condition (Definition 5.6.1): x_0 appears neither in χ, nor in $\exists x\, \varphi$, nor in a premise, nor in an assumption of a box still open at line k.

What the hypothesis gives. Line h is available at line k, so $\Gamma_h \subseteq \Gamma_k$ (Lemma 4.7.1 (a)), and $S(h)$ holds: every structure interpreting every symbol of the derivation, and every assignment satisfying Γ_h, satisfy $\exists x\, \varphi$. The box is cited by a box rule, so its last line has context $\Gamma_k, \varphi[x_0/x]$ (Lemma 4.7.1 (b), extended in the subsection "Contexts, once more"),

and $S(b)$ says that every such structure and assignment satisfying Γ_k and $\varphi[x_0/x]$ satisfy χ.

The argument. Let $\mathfrak{M}$ be a structure interpreting every symbol of the derivation (as in Step 2), with domain D, and ℓ an assignment with $\mathfrak{M}, \ell \models \Gamma_k$. Then $\mathfrak{M}, \ell \models \Gamma_h$, so $\mathfrak{M}, \ell \models \exists x\, \varphi$; by (S7), choose $d \in D$ with $\mathfrak{M}, \ell[x \mapsto d] \models \varphi$. Let $\ell' := \ell[x_0 \mapsto d]$.

1. $\mathfrak{M}, \ell' \models \Gamma_k$. The formulas of Γ_k are the premises and the assumptions of the boxes open at line k; by the side condition, x_0 appears in none of them. ℓ' differs from ℓ only at x_0, so Lemma 7.1.1 applies to each formula of Γ_k.
2. $\mathfrak{M}, \ell' \models \varphi[x_0/x]$. Since x_0 does not appear in $\exists x\, \varphi$, it is not x and does not appear in φ. So $\ell'[x \mapsto d]$ and $\ell[x \mapsto d]$ differ at most at x_0, which is not in φ, and $\mathfrak{M}, \ell'[x \mapsto d] \models \varphi$ by Lemma 7.1.1. The term x_0 is free for x in φ, as no quantifier of φ is on x_0, and $x_0^{\mathfrak{M},\ell'} = d$; so Lemma 7.3.1 gives $\mathfrak{M}, \ell' \models \varphi[x_0/x]$.
3. By 1, 2 and the hypothesis for the box, $\mathfrak{M}, \ell' \models \chi$.
4. ℓ' and ℓ differ only at x_0, which does not appear in χ; by Lemma 7.1.1, $\mathfrak{M}, \ell \models \chi$.

Since $\mathfrak{M}$ and ℓ were arbitrary, this proves $S(k)$.

Marking scheme. (10 points)

- 2: states the data and the full side condition, and the context $\Gamma_k, \varphi[x_0/x]$ of the box's last line.
- 2: obtains a witness d from the cited line through the induction hypothesis and (S7), and sets x_0 to d.
- 2: shows the new assignment still satisfies Γ_k, using "x_0 in no premise and no open assumption".
- 2: shows the new assignment satisfies $\varphi[x_0/x]$, using "x_0 not in $\exists x\, \varphi$" and the substitution lemma.
- 2: applies the hypothesis for the box, and returns to ℓ using "x_0 not in χ".

A proof that treats the box's assumption and the context in the other order earns the same points. Fatal errors (0 for the step): setting x_0 to an arbitrary element rather than the witness; concluding $\mathfrak{M}, \ell \models \chi$ directly from $\mathfrak{M}, \ell' \models \chi$ without the condition on χ.

Exercise. C2: "Free for" is needed

Show that the side condition "t is free for x in φ" of ∀E cannot be dropped. Give a two-line "derivation" that uses ∀E with a term that is not free for x, a structure in which its premise is true and its conclusion false, and the step of the proof of Lemma 7.3.1 that fails.

Solution (Hint 1). Look for a premise $\forall x\, \varphi$ in which φ contains a quantifier $\forall y$ and x occurs free inside its scope; instantiate at $t = y$.

Solution (Hint 2). Try $\varphi = \neg\forall y\, (y \leq x)$, "$x$ is not an upper bound of everything", in a structure with no largest element.

Solution. Take $\varphi = \neg\forall y\,(y \leq x)$ and $t = y$. The free occurrence of x in φ lies in the scope of $\forall y$, and y occurs in t, so t is **not** free for x in φ. Ignoring this, $\forall$E would allow:

$$\begin{array}{lll} 1 & \forall x\,\neg\forall y\,(y \leq x) & \\ \hline 2 & \neg\forall y\,(y \leq y) & \forall\text{E } 1 \end{array}$$

In $(\mathbb{Z}, \leq)$, the premise says "for every integer x, not every integer y is $\leq x$", that is, no integer is largest: for each x, the integer $x + 1$ is not $\leq x$. It is true. The conclusion says "not every integer y has $y \leq y$", which is false, since $\leq$ is reflexive. So the premise is true and the conclusion false, and a rule allowing this step would be unsound.

The step of Lemma 7.3.1 that fails is the second quantifier case, applied to $\forall y\,(y \leq x)$: it needs y not in t, so that the value of t does not change when (S6) sets y to each d. Here $t = y$, and its value is the new d each time: $\forall y\,(y \leq y)$ compares each d with itself, while $\forall y\,(y \leq x)$ at the value a of t compares each d with the fixed a. The legal route renames first (Section 5.2): $\forall x\,\neg\forall z\,(z \leq x)$, and then $\forall$E at y gives $\neg\forall z\,(z \leq y)$, which is true in $(\mathbb{Z}, \leq)$ under every assignment.

Marking scheme. (10 points)

- 3: a premise and a $\forall$E step at a term that is not free for x, with the capture identified.
- 4: a structure, with the premise checked true and the conclusion checked false.
- 3: names the quantifier case of Lemma 7.3.1 and the fact it needs (y not in t).

Fatal errors (0 for the step): a "counter-model" in which the premise is false; a step in which t is in fact free for x.

7.4 Theories and Their Models

Section 7.2 called a structure in which three sentences are true a model of them. Mathematics works this way all the time: a list of axioms describes a kind of structure, and a theorem about that kind is whatever follows from the axioms. This section makes that precise, and answers the chapter's first question: the axioms of a partial order do not prove that any two elements are comparable.

The plan has five steps. First, theories and their models are defined. A gallery follows: equivalence relations, orders, groups and fields, each written as axioms, each with a model and a non-model. Then one theorem is derived from the group axioms, to show what a theorem of a theory is. Then comes consistency, which a single model proves. Last comes independence, which two models prove.

7.4.1 Theories and their models

Fix the symbols (constants, function and predicate symbols) that the axioms will use, as in Definition 7.1.1.

A theory is a collection of sentences taken as axioms; its theorems are what can be derived from finitely many of them.

Definition 7.4.1. Theory

A **theory** T is a set of sentences in the chosen symbols, called its **axioms**; the set may be finite or infinite. A sentence σ is **derivable from** T, written $T \vdash \sigma$, if $\varphi_1, \dots, \varphi_n \vdash \sigma$ for some finite list $\varphi_1, \dots, \varphi_n$ of axioms of T ($n \geq 0$); such a σ is a **theorem** of T.

In words: a derivation is finite, so it uses finitely many axioms, even when T has infinitely many. So "$T \vdash \sigma$" reads Definition 1.2.3 through a finite list, and needs nothing new. The case $n = 0$ says that a sentence derivable from no premises at all is a theorem of every theory. At the other end, each axiom is a theorem of its own theory: the derivation whose only line is the axiom, as a premise, has it as its last line. The theorems are everything in between: what the axioms, used finitely often, force.

A model of a theory is a structure in which every axiom is true.

Definition 7.4.2. Model of a Theory

A structure $\mathfrak{M}$ for the symbols of T is a **model of** T, written $\mathfrak{M} \models T$, if every axiom of T is true in $\mathfrak{M}$. A sentence σ is a **semantic consequence** of T, written $T \models \sigma$, if σ is true in every model of T.

The axioms are sentences, so no assignment is needed: by Lemma 7.1.1 each axiom is true or false in $\mathfrak{M}$ outright. And "every model of T" is the theory's version of "every structure satisfying Γ" in Definition 7.1.4.

Quick check. Let T be the three sentences of the strict order with no largest element (Section 7.2). Is $(\{0\}, <)$ a model of T? Is the set $\mathbb{Q}_{>0}$ of positive rationals with $<$?

Solution. Not the first: $\forall x\, \neg(x < x)$ and transitivity hold, but $\forall x\, \exists y\, (x < y)$ fails at 0, since nothing is above it. The second is a model: no rational is below itself, $<$ is transitive, and for each positive d the witness $d + 1$ is again positive and above d.

7.4.2 A gallery of theories

Each theory below is given by its axioms, written as sentences and named, with a model and a non-model. A non-model is checked by naming the axiom that fails and where. The names in parentheses are local to this chapter, except the field clauses (F1)–(F9), which keep their series-wide numbering; none of them is one of the clauses (S1)–(S7) of Definition 7.1.3.

Example. Axioms, Models and Non-models

Equivalence relations. One two-place predicate symbol R, and three axioms:

$$\begin{aligned} &\text{(Refl)} && \forall x\, R(x, x), \\ &\text{(Sym)} && \forall x\, \forall y\, (R(x, y) \to R(y, x)), \\ &\text{(Trans)} && \forall x\, \forall y\, \forall z\, ((R(x, y) \wedge R(y, z)) \to R(x, z)). \end{aligned}$$

A model: $D = \{a, b, c\}$, with $R^{\mathfrak{M}}$ relating a and b both ways and every element to itself (Figure A, left). Each loop gives (Refl); the only pairs of different elements are (a, b) and (b, a), which give (Sym); and every chain of two arrows ends where a loop or one of these pairs already is, which gives (Trans). In numbers: "x and y have the same parity" on $\mathbb{Z}$. A non-model: on $\mathbb{Z}$, read $R(x, y)$ as $|x - y| \leq 1$. It is reflexive and symmetric, and (Trans) fails at $0, 1, 2$: $|0 - 1| \leq 1$ and $|1 - 2| \leq 1$, but $|0 - 2| = 2$.

Partial and linear orders. One two-place symbol $\leq$. The axioms (Refl) and (Trans), written with $\leq$, together with

$$\text{(Antisym)} \quad \forall x \, \forall y \, \big((x \leq y \wedge y \leq x) \to x = y\big),$$

make the theory of a **partial order**. Adding

$$\text{(Conn)} \quad \forall x \, \forall y \, (x \leq y \vee y \leq x)$$

gives the theory of a **linear order**. A model of the first and not the second: divisibility on $\{1, 2, 3, 6\}$ (the right panel of Figure F in Section 7.1), where 2 and 3 are incomparable, so (Conn) fails at $(2, 3)$. A model of both: $(\mathbb{N}, \leq)$. Smallest of all, the **antichain** $D = \{a, b\}$ with $\leq^{\mathfrak{M}} = \{(a, a), (b, b)\}$ (Figure A, right) is a partial order: the loops give (Refl), and (Antisym) and (Trans) only ever meet pairs (d, d). (Conn) fails at (a, b).

Strict and dense orders. One two-place symbol $<$. The axioms

$$\begin{aligned} &\text{(Irrefl)} \quad \forall x \, \neg(x < x), \\ &\text{(Trans)} \quad \forall x \, \forall y \, \forall z \, \big((x < y \wedge y < z) \to x < z\big) \end{aligned}$$

make a **strict order**; with

$$\text{(Tri)} \quad \forall x \, \forall y \, (x < y \vee x = y \vee y < x)$$

a **strict linear order**; and with

$$\text{(Dense)} \quad \forall x \, \forall y \, \big(x < y \to \exists z \, (x < z \wedge z < y)\big)$$

a **dense** one. A model of all four: $(\mathbb{Q}, <)$, where $(p + q)/2$ lies between $p < q$. A non-model of (Dense): $(\mathbb{Z}, <)$, at $0 < 1$, since no integer lies strictly between.

Groups. A two-place function symbol $\cdot$ and a constant e, with

$$\begin{aligned} &\text{(Assoc)} \quad \forall x \, \forall y \, \forall z \, \big((x \cdot y) \cdot z = x \cdot (y \cdot z)\big), \\ &\text{(Id)} \quad \forall x \, (x \cdot e = x \wedge e \cdot x = x), \\ &\text{(Inv)} \quad \forall x \, \exists y \, (x \cdot y = e \wedge y \cdot x = e). \end{aligned}$$

A model: $(\mathbb{Z}; +, 0)$, with $\cdot$ read as $+$ and e as 0; the inverse of d is $-d$. A non-model: $(\mathbb{N}; +, 0)$, where (Inv) fails at 1, since no natural number added to 1 gives 0. Commutativity,

$$\text{(Comm)} \quad \forall x \, \forall y \, (x \cdot y = y \cdot x),$$

is **not** an axiom; the next example shows why.

Fields. Constants 0, 1 and two-place function symbols $+$, $\cdot$, with ten axioms, the nine (F1)–(F9), numbered as in the field axioms used throughout this series, and $(1 \neq 0)$:

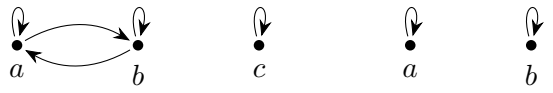
(F1)	$\forall x\,\forall y\,(x + y = y + x)$,	(F2)	$\forall x\,\forall y\,\forall z\,\big((x + y) + z = x + (y + z)\big)$,
(F3)	$\forall x\,(x + 0 = x)$,	(F4)	$\forall x\,\exists y\,(x + y = 0)$,
(F5)	$\forall x\,\forall y\,\forall z\,\big((x \cdot y) \cdot z = x \cdot (y \cdot z)\big)$,	(F6)	$\forall x\,\forall y\,(x \cdot y = y \cdot x)$,
(F7)	$\forall x\,(x \cdot 1 = x)$,	(F8)	$\forall x\,\big(\neg(x = 0) \to \exists y\,(x \cdot y = 1)\big)$,
(F9)	$\forall x\,\forall y\,\forall z\,(x \cdot (y + z) = x \cdot y + x \cdot z)$,	$(1 \neq 0)$	$\neg(0 = 1)$.

Models: $\mathbb{Q}$ and $\mathbb{R}$, with the usual operations, and the **two-element field**, with domain $\{0, 1\}$ and the tables

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

(the sum of two equal elements is 0; the products are the usual ones). A non-model: $\mathbb{Z}$, where (F8) fails at 2: no integer y has $2y = 1$.



an equivalence relation the antichain: a partial order, not linear

Figure A. Left: an equivalence relation on $\{a, b, c\}$, whose classes are $\{a, b\}$ and $\{c\}$. Right: the antichain on $\{a, b\}$.

The non-models are as instructive as the models: a non-model with every axiom but one true shows exactly what that one axiom adds.

Quick check. In the two-element field, which element is the additive inverse of 1, and which is the reciprocal of 1? Why does the field axiom for reciprocals ask nothing of 0?

Solution. 1 in both cases: $1 + 1 = 0$ and $1 \cdot 1 = 1$. The axiom (F8) reads $\forall x\,\big(\neg(x = 0) \to \exists y\,(x \cdot y = 1)\big)$, and at 0 its hypothesis fails, so the implication holds by (S5).

The groups part promised a model of the group axioms in which (Comm) is false. Every group of numbers under addition or multiplication is commutative, so the model comes from geometry.

Example. The Six Symmetries of a Triangle

Label the corners of an equilateral triangle 1, 2, 3 (Figure T). A **symmetry** is a motion that puts the triangle back on itself, recorded by where it sends each corner. There are six:

- e: nothing moves;
- r: a turn by a third, $1 \mapsto 2$, $2 \mapsto 3$, $3 \mapsto 1$;
- q: a turn by two thirds, $1 \mapsto 3$, $2 \mapsto 1$, $3 \mapsto 2$;
- the flips a (fixing 1, swapping 2 and 3), b (fixing 2) and c (fixing 3).

Read $x \cdot y$ as "**first** x, **then** y". For instance $a \cdot r$ sends 1 to 1 and then to 2; 2 to 3 and then

to 1; 3 to 2 and then to 3. It fixes 3 and swaps 1, 2, so $a \cdot r = c$. Computed in the same way, the whole table, with row x, column y and entry $x \cdot y$, is

$\cdot$	e	r	q	a	b	c
e	e	r	q	a	b	c
r	r	q	e	b	c	a
q	q	e	r	c	a	b
a	a	c	b	e	q	r
b	b	a	c	r	e	q
c	c	b	a	q	r	e

Take the domain $\{e, r, q, a, b, c\}$, the table as the meaning of $\cdot$, and e as the meaning of the constant e. Check the group axioms.

- (Id): the e row and the e column repeat the column and row heads.
- (Inv): each row contains e, and in the same place as its column does. So $r \cdot q = e = q \cdot r$, and each flip undoes itself: $a \cdot a = b \cdot b = c \cdot c = e$.
- (Assoc): "first x, then y, then z" is one sequence of motions, however it is bracketed. The fold below says this once more, carefully.

Now $a \cdot r = c$ and $r \cdot a = b$, so (Comm) fails at (a, r). So this is a model of the group axioms in which (Comm) is false.

Two elements would not do: in a group $\{e, g\}$, every product with e is fixed by (Id), and $g \cdot g$ commutes with itself, so (Comm) holds.

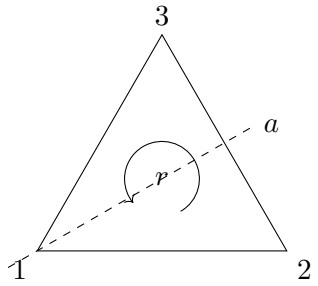


Figure T. The corners 1, 2, 3; the turn r about the center, carrying 1 to 2, 2 to 3 and 3 to 1; and the line of the flip a, through corner 1 and the midpoint of the opposite side.

↻ **Second reading** · *Why the table is associative without 216 checks*

Checking (Assoc) entry by entry would mean $6 \cdot 6 \cdot 6 = 216$ triples (x, y, z), each needing four look-ups in the table. There is a shorter way. A symmetry is a rearrangement of the corners, and $x \cdot y$ is the rearrangement that applies x and then y to each corner. Follow one corner i through $(x \cdot y) \cdot z$: first $x \cdot y$ sends it to where y sends the place x sent it, and then z moves it on. Through $x \cdot (y \cdot z)$: first x moves it, and then $y \cdot z$ applies y and then z. Either way, corner i goes through x, then y, then z, and lands in the same place. The two products send every corner to the same corner, so they are the same symmetry. The table only records the results, so it is associative because composing rearrangements is.

Quick check. From the table, what is $a \cdot a$? What is $r \cdot r \cdot r$, bracketed either way?

Solution. $a \cdot a = e$. And $(r \cdot r) \cdot r = q \cdot r = e$, while $r \cdot (r \cdot r) = r \cdot q = e$: three turns by a third bring every corner home.

7.4.3 A theorem of a theory

A theorem of a theory is derived from finitely many of its axioms with the rules of Chapters 1, 2 and 5; nothing else is used. Section 5.4 has already derived one, in Exercise (B5: Derive or break: properties of a relation, used later) (c): from (Irrefl) and (Trans), a strict order is asymmetric. Here is a theorem of the group axioms.

Proposition 7.4.1. A Group Has Only One Identity

From (Id), the sentence

$$\forall x\,(\forall y\,(y \cdot x = y \wedge x \cdot y = y) \to x = e)$$

is derivable: an element that acts as an identity on every element is e.

Proof. The plan: fix x_0 and assume that it acts as an identity; apply that assumption at e, and (Id) at x_0. The two give two values of the one product $x_0 \cdot e$, and =E equates them.

1	$\forall x\,(x \cdot e = x \wedge e \cdot x = x)$		
2	x_0	$\forall y\,(y \cdot x_0 = y \wedge x_0 \cdot y = y)$	asm, for $\to$I
3		$e \cdot x_0 = e \wedge x_0 \cdot e = e$	$\forall$E 2
4		$x_0 \cdot e = e$	$\wedge$E 3
5		$x_0 \cdot e = x_0 \wedge e \cdot x_0 = x_0$	$\forall$E 1
6		$x_0 \cdot e = x_0$	$\wedge$E 5
7		$x_0 = e$	=E 6, 4
8	$\forall y\,(y \cdot x_0 = y \wedge x_0 \cdot y = y) \to x_0 = e$		$\to$I 2–7
9	$\forall x\,(\forall y\,(y \cdot x = y \wedge x \cdot y = y) \to x = e)$		$\forall$I 2–8

Line 7 is =E with φ the formula $z = e$: line 4 is $\varphi[x_0 \cdot e/z]$, and line 6 says $x_0 \cdot e = x_0$, so the rule gives $\varphi[x_0/z]$. At line 9, x_0 appears in no premise, in no open assumption, and not in line 9.

■

So in **every** group the identity is the only element that acts as one, by the soundness statement of Section 7.2, without checking any group. In $(\mathbb{Z}; +, 0)$, 0 is the only integer x with $d + x = d$ and $x + d = d$ for every d; in the triangle group, only the e row and the e column repeat the

heads. Only (Id) was used. A theorem of a theory says something about all its models at once; that is what a theory is for.

7.4.4 Consistency

Section 3.2 called a list of formulas inconsistent when it derives $\bot$ (Definition 3.2.1). A theory may be infinite, so its version goes through finite lists.

Definition 7.4.3. Inconsistent Theory

A theory T is **inconsistent** if $T \vdash \bot$, that is, if $\varphi_1, \ldots, \varphi_n \vdash \bot$ for some finite list of axioms of T; otherwise T is **consistent**.

An inconsistent theory derives every sentence, since $\bot$E turns $\bot$ into anything. So it says nothing, and consistency is the least one asks of axioms. Deriving $\bot$ shows inconsistency: one derivation is enough. Consistency is harder to show by derivations, since it says that **no** derivation from the axioms ends with $\bot$, and there are infinitely many derivations to rule out. Section 7.2 met the same difficulty with non-derivability, and solved it with a counter-model. A model does the same here: it rules out every derivation of $\bot$ at once.

Proposition 7.4.2. A Theory with a Model Is Consistent

If some structure is a model of T, then T is consistent.

Proof. Let $\mathfrak{M} \models T$, and suppose $\varphi_1, \ldots, \varphi_n \vdash \bot$ for axioms φ_i of T. By the soundness statement of Section 7.2 (proved in the optional Section 7.3), $\varphi_1, \ldots, \varphi_n \models \bot$. Each φ_i is true in $\mathfrak{M}$, so $\mathfrak{M}, \ell \models \bot$ for any assignment ℓ, which (S3) forbids. So no such list exists. ∎

So each model in the gallery proves its theory consistent. The converse, that every consistent theory has a model, is also true; it is the completeness theorem for predicate logic, stated without proof in the optional Section 7.6. *Consistent* is a word about derivations, and *has a model* (satisfiable) is one about structures. The two are never swapped, and where they agree the text says which theorem makes them agree.

The same reasoning shows that a sentence is **not** a theorem. To show $T \nvdash \sigma$, make the counterexample set $T, \neg\sigma$ true (Definition 1.1.3): find a model of T in which σ is false.

Warning. A model in which σ is **true** shows nothing about $T \vdash \sigma$. Comparability, (Conn), is true in $(\mathbb{N}, \leq)$, a model of the partial order axioms, and it is not a theorem of them, as the next subsection shows.

Quick check. Is the theory $\{\exists x\, P(x),\ \forall x\, \neg P(x)\}$ consistent?

Solution. No. Open an ∃E box on the first axiom, with witness x_0 and assumption $P(x_0)$; ∀E on the second gives $\neg P(x_0)$, and ¬E gives $\bot$, which ∃E carries out of the box, since x_0 is not in $\bot$, in the first axiom or in a premise. (Equally, no structure is a model; but that alone would not show inconsistency without the completeness theorem.)

7.4.5 Independence

One model with σ false shows that σ is not a theorem. Two models, one each way, show more.

Proposition 7.4.3. Independence by Models

Let T be a theory and σ a sentence.

(a) If some model of T makes σ false, then $T \nvdash \sigma$.

(b) If, besides, some model of T makes σ true, then $T \nvdash \neg\sigma$ as well; σ is then **independent** of T: neither it nor its negation is a theorem.

Proof.

(a) If $\varphi_1, \ldots, \varphi_n \vdash \sigma$ for axioms of T, then $\varphi_1, \ldots, \varphi_n \models \sigma$ by the soundness statement of Section 7.2 (proved in the optional Section 7.3). The model satisfies every φ_i, so it satisfies σ, against the hypothesis. (b) Apply (a) to $\neg\sigma$, which the second model makes false, by (S4).

■

Now the chapter's question. The antichain of Figure A is a model of the partial order axioms in which (Conn) is false, and $(\mathbb{N}, \leq)$ is one in which it is true. So comparability is independent of the axioms of a partial order: they neither prove nor refute it. That is why "linear order" needs (Conn) as an axiom of its own.

The same goes for groups. $(\mathbb{Z}; +, 0)$ and the triangle group are models of the group axioms; (Comm) is true in the first and false in the second. So commutativity is independent of the group axioms. Mathematicians call the groups that satisfy it **abelian**, and study the others too.

Finding the two models. The model in which σ is true is usually the one the axioms were written for, such as $(\mathbb{N}, \leq)$ or $(\mathbb{Z}; +, 0)$. The model in which σ is false is a counter-model in the sense of Section 7.2, and it is found by the same backward reading: write down what $\neg\sigma$ demands, then what each axiom demands, and build the smallest structure meeting both. For (Conn) the demand is two incomparable elements, and the axioms of a partial order then ask only for a loop at each; that is the antichain.

Independence also tells whether an axiom list has a spare member: an axiom that follows from the others could be dropped from the list without changing the theorems. The three axioms of an equivalence relation have none.

Example. Reflexive, Symmetric, Transitive: Each Independent of the Other Two

Three structures for one two-place symbol R (Figure I). The first two are the counter-readings

(d) and (f) of Exercise (B5: Derive or break: properties of a relation, used later), now as models, with a, b in place of 0, 1.

- *Not reflexive*: $D = \{a, b\}$, $R^{\mathfrak{M}} = \{(a, a)\}$. Symmetric and transitive, since the only pair is (a, a); (Refl) fails at b.
- *Not symmetric*: $D = \{a, b\}$, $R^{\mathfrak{M}} = \{(a, a), (b, b), (a, b)\}$. Reflexive; transitive, since every chain of two arrows has a loop as one of them; (Sym) fails at (a, b), since (b, a) is missing.
- *Not transitive*: $D = \{a, b, c\}$, $R^{\mathfrak{M}}$ the loops at a, b, c together with (a, b), (b, a), (b, c), (c, b). Reflexive; symmetric, since the arrows between different elements come in pairs; (Trans) fails at a, b, c, since (a, b) and (b, c) are in $R^{\mathfrak{M}}$ and (a, c) is not.

By Proposition 7.4.3 (a), no one of the three axioms follows from the other two.

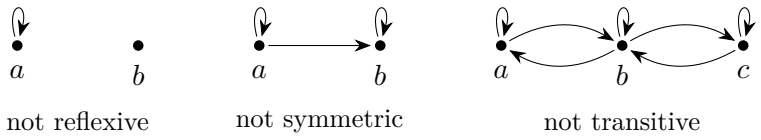


Figure I. Three relations, each with exactly two of the three properties of an equivalence relation.

Warning. A tempting "proof" that a symmetric and transitive relation is reflexive: "$R(x, y)$ gives $R(y, x)$ by symmetry, and the two give $R(x, x)$ by transitivity." The step from nothing to $R(x, y)$ is the error: it needs some y related to x. In the first structure above, b is related to nothing, and $R(b, b)$ fails. Section 5.4 broke the claim with this reading, Exercise (B5: Derive or break: properties of a relation, used later) (d), and here it is a model of (Sym) and (Trans) in which (Refl) is false.

Quick check. Show that (Dense) is independent of the strict linear order axioms.

Solution. $(\mathbb{Q}, <)$ is a strict linear order in which (Dense) holds; $(\mathbb{Z}, <)$ is one in which it fails, at $0 < 1$. Apply Proposition 7.4.3.

Remark (Looking ahead). *Sets, Relations, and Cardinality*, Chapter 2 studies equivalence relations and partial orders as relations on a set, with their maximal and greatest elements; *Algebraic Preliminaries*, Chapters 2 and 3 take fields and groups as theories and study their models.

7.4.6 Exercises

7.4.6.1 A. Check your understanding

Exercise. A1

(a) Define a theory, a model of a theory, and $T \vdash \sigma$.

(b) Why does Definition 7.4.3 say "finite list of axioms"?

(c) In the proof of Proposition 7.4.3 (a), where is the soundness statement used? Why does a model in which σ is true not show $T \vdash \sigma$?

(d) What two models show σ independent of T?

(e) True or false: every theory with a finite model is consistent.

Solution. *(a)* A theory is a set of sentences, its axioms (Definition 7.4.1). A model of T is a structure in which every axiom of T is true (Definition 7.4.2). $T \vdash \sigma$ means $\varphi_1, \dots, \varphi_n \vdash \sigma$ for some finite list of axioms of T.

(b) A derivation has finitely many lines, so it uses finitely many premises, and a theory may be infinite.

(c) To pass from $\varphi_1, \dots, \varphi_n \vdash \sigma$ to $\varphi_1, \dots, \varphi_n \models \sigma$. A model in which σ is true is compatible with $T \vdash \sigma$ and with $T \nvdash \sigma$ alike: comparability is true in $(\mathbb{N}, \leq)$ and is not a theorem of the partial order axioms.

(d) A model of T in which σ is false, and one in which it is true.

(e) True, by Proposition 7.4.2: a finite model is a model.

Exercise. A2: Drill

Each part takes a line or two. The parts run from reading an axiom, through deciding axioms in a given structure, to judging claims about theories and building a model.

(a) Write "R is symmetric" as a sentence.

(b) Which axiom of a partial order fails in $(\mathbb{Z}, <)$, with $\leq$ read as $<$?

(c) In the triangle table, what are $q \cdot a$ and $a \cdot q$?

(d) Is $D = \{a, b\}$, $R^{\mathfrak{M}} = \{(a, b), (b, a)\}$ a model of the equivalence relation axioms? Which of the three hold?

(e) Is $(\mathbb{Q} \setminus \{0\}; \cdot, 1)$, with the group's $\cdot$ read as multiplication and e as 1, a model of the group axioms?

(f) On $D = \{a\}$, which relations make a model of the strict order axioms? Of the partial order axioms?

(g) On $D = \{a, b\}$ with $(a, b) \in R^{\mathfrak{M}}$, add the fewest pairs that make $R^{\mathfrak{M}}$ an equivalence relation.

(h) Give a model of the dense strict linear order axioms other than $(\mathbb{Q}, <)$.

(i) True or false: if σ is false in some structure, then $T \nvdash \sigma$.

(j) Is the empty theory consistent?

(k) Is the theory with the two axioms $\forall x\, \neg R(x, x)$ and $\forall x\, \forall y\, R(x, y)$ consistent?

(l) Give a model of the partial order axioms with three elements in which exactly one pair of different elements is comparable.

Solution.

(a) $\forall x \forall y\,(R(x,y) \to R(y,x))$: whenever x is related to y, y is related to x.

(b) (Refl): $0 < 0$ is false.

(c) $q \cdot a = c$ and $a \cdot q = b$.

(d) No. (Sym) holds; (Refl) fails, since there are no loops; (Trans) fails, since (a,b) and (b,a) are in $R^{\mathfrak{M}}$ and (a,a) is not.

(e) Yes: multiplication is associative, 1 is an identity, and each d has the inverse $1/d$, again a non-zero rational.

(f) Strict order: exactly $<^{\mathfrak{M}} = \emptyset$. The loop (a,a) breaks (Irrefl), and with no pairs (Trans) has no chain to check. Partial order: exactly $\leq^{\mathfrak{M}} = \{(a,a)\}$; with $\emptyset$, (Refl) fails at a.

(g) (a,a), (b,b) and (b,a).

(h) $(\mathbb{R}, <)$, or the rationals strictly between 0 and 1 with $<$.

(i) False: the structure must be a **model of** T. With $T = \{\forall x\, R(x,x)\}$ and σ the same sentence, σ is false in every structure with $R^{\mathfrak{M}} = \emptyset$, yet $T \vdash \sigma$ in one line.

(j) Yes: every structure is a model of it, so Proposition 7.4.2 applies.

(k) No: $\forall$E twice on the second axiom gives $R(x_0,x_0)$, $\forall$E on the first gives $\neg R(x_0,x_0)$, and $\neg$E gives $\bot$. (Any variable or constant may be the term.)

(l) $D = \{a,b,c\}$, $\leq^{\mathfrak{M}}$ the three loops and (a,b).

7.4.6.2 B. Practice

Exercise. B1: Construct a relation, or show none exists, used later

On $D = \{1,2,3,4\}$, give a relation that is:

(i) reflexive and transitive, but neither symmetric nor antisymmetric;

(ii) symmetric and transitive, not reflexive, with 1, 2 and 3 each related to something;

(iii) irreflexive, symmetric, transitive and non-empty; or show that no domain has one.

(iv) Which of your answers refutes the "proof" in the last warning of this section, and at which element?

Solution.

(i) R = the four loops together with $(1,2)$, $(2,1)$, $(1,3)$, $(2,3)$. Reflexive: the loops. Transitive: the chains $(x,y),(y,z)$ of two non-loop pairs ($x \neq y$ and $y \neq z$) are $(1,2),(2,1)$, giving $(1,1)$; $(2,1),(1,2)$, giving $(2,2)$; $(1,2),(2,3)$, giving $(1,3)$; and $(2,1),(1,3)$, giving $(2,3)$; all in R, and a chain through a loop gives back its other pair. Not symmetric: $(1,3)$ without $(3,1)$. Not antisymmetric: $(1,2)$ and $(2,1)$ with $1 \neq 2$.

(ii) $R = \{(1,1),(1,2),(2,1),(2,2),(3,3)\}$: symmetric and transitive (the classes $\{1,2\}$ and $\{3\}$), not reflexive at 4.

(iii) None, on any domain. If $(d,e) \in R$, symmetry gives (e,d), and transitivity applied to $(d,e),(e,d)$ gives (d,d), against irreflexivity.

(iv) (ii), at 4, which is related to nothing, so the "proof" has no y to start from.

Used later: Section 7.5 comes back to the relation of (ii).

Marking scheme. (10 points) - 3: (i), a relation with all four properties checked, transitivity by listing the chains. - 2: (ii). - 3: (iii), the argument from an arbitrary pair. - 2: (iv). Fatal error (0 for (iii)): an argument that assumes every element is related to something.

Exercise. B2: A sentence true in one structure and false in the other

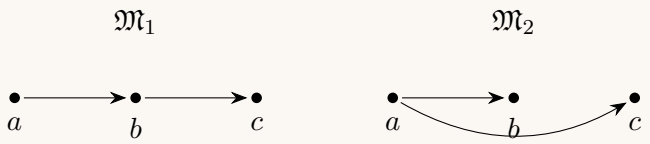
For each pair, find a sentence true in the first structure and false in the second, and read it back in words.

(a) $(\mathbb{Z}; +, 0)$ and $(\mathbb{N}; +, 0)$.

(b) $(\mathbb{Q}; \cdot, 0, 1)$ and $(\mathbb{Z}; \cdot, 0, 1)$.

(c) $\mathfrak{M}_1$: $D = \{a, b, c\}$, $R^{\mathfrak{M}_1} = \{(a, b), (b, c)\}$; and $\mathfrak{M}_2$: $D = \{a, b, c\}$, $R^{\mathfrak{M}_2} = \{(a, b), (a, c)\}$ (drawn below).

(d) $\mathfrak{M}_2$ and $\mathfrak{M}_1$, in that order.



Solution. *(a)* $\forall x\, \exists y\, (x + y = 0)$, "every element has an additive inverse": true in $\mathbb{Z}$, with $y = -x$; false in $\mathbb{N}$ at 1.

(b) $\forall x\, (\neg(x = 0) \to \exists y\, (x \cdot y = 1))$, "every non-zero element has a reciprocal": true in $\mathbb{Q}$, with $y = 1/x$; false in $\mathbb{Z}$ at 2.

(c) $\exists x\, \exists y\, \exists z\, (R(x, y) \wedge R(y, z))$, "there is a path of two arrows": in $\mathfrak{M}_1$, a, b, c; in $\mathfrak{M}_2$ no arrow leaves b or c, so no path of two arrows exists.

(d) $\exists x\, \exists y\, \exists z\, (R(x, y) \wedge R(x, z) \wedge \neg(y = z))$, "some element has arrows to two different elements": in $\mathfrak{M}_2$, a, b, c; in $\mathfrak{M}_1$ each element has at most one outgoing arrow.

Any other correct sentence earns full marks.

Marking scheme. (10 points) - 2.5 per part: 1 for a sentence true in the first structure (checked), 1 for false in the second (with the failing element), 0.5 for the reading back. Fatal error (0 for the part): a formula with a free variable.

Exercise. B3: Derive a theorem of a theory

From the three group axioms, derive $\forall x\, (x \cdot x = x \to x = e)$: the only element equal to its own square is the identity. Which axioms did you use? Show by a model that (Inv) cannot be left out.

Solution. *Plan.* Fix x_0 and assume $x_0 \cdot x_0 = x_0$. Take an inverse y_0 by $\exists$E, multiply the assumption by y_0 on the left, and regroup: in words, $e = y_0 \cdot x_0 = y_0 \cdot (x_0 \cdot x_0) = (y_0 \cdot x_0) \cdot x_0 = e \cdot x_0 = x_0$.

1		$\forall x\,\forall y\,\forall z\,((x \cdot y) \cdot z = x \cdot (y \cdot z))$	
2		$\forall x\,(x \cdot e = x \wedge e \cdot x = x)$	
3		$\forall x\,\exists y\,(x \cdot y = e \wedge y \cdot x = e)$	
4	x_0	$x_0 \cdot x_0 = x_0$	asm, for $\to$I
5		$\exists y\,(x_0 \cdot y = e \wedge y \cdot x_0 = e)$	$\forall$E 3
6		y_0 | $x_0 \cdot y_0 = e \wedge y_0 \cdot x_0 = e$	asm, for $\exists$E
7		$y_0 \cdot x_0 = e$	$\wedge$E 6
8		$\forall y\,\forall z\,((y_0 \cdot y) \cdot z = y_0 \cdot (y \cdot z))$	$\forall$E 1
9		$\forall z\,((y_0 \cdot x_0) \cdot z = y_0 \cdot (x_0 \cdot z))$	$\forall$E 8
10		$(y_0 \cdot x_0) \cdot x_0 = y_0 \cdot (x_0 \cdot x_0)$	$\forall$E 9
11		$(y_0 \cdot x_0) \cdot x_0 = y_0 \cdot x_0$	=E 4, 10
12		$e \cdot x_0 = y_0 \cdot x_0$	=E 7, 11
13		$e \cdot x_0 = e$	=E 7, 12
14		$x_0 \cdot e = x_0 \wedge e \cdot x_0 = x_0$	$\forall$E 2
15		$e \cdot x_0 = x_0$	$\wedge$E 14
16		$x_0 = e$	=E 15, 13
17		$x_0 = e$	$\exists$E 5, 6–16
18		$x_0 \cdot x_0 = x_0 \to x_0 = e$	$\to$I 4–17
19		$\forall x\,(x \cdot x = x \to x = e)$	$\forall$I 4–18

The four =E steps, each with its φ: line 11 with φ the formula $(y_0 \cdot x_0) \cdot x_0 = y_0 \cdot w$, line 10 being $\varphi[x_0 \cdot x_0/w]$; line 12 with $w \cdot x_0 = y_0 \cdot x_0$, line 11 being $\varphi[y_0 \cdot x_0/w]$; line 13 with $e \cdot x_0 = w$, line 12 being $\varphi[y_0 \cdot x_0/w]$; line 16 with $w = e$, line 13 being $\varphi[e \cdot x_0/w]$. At line 17, y_0 is not in $x_0 = e$, not in line 5, in no premise and in no open assumption. At line 19, x_0 is in no premise, in no assumption open outside its box, and not in line 19.

All three axioms are used. (Inv) cannot be left out: $(\{0,1\}; \cdot, 1)$, with the usual product and e read as 1, satisfies (Assoc), since multiplication of numbers is associative, and (Id), since $0 \cdot 1 = 1 \cdot 0 = 0$ and $1 \cdot 1 = 1$. But $0 \cdot 0 = 0$ and $0 \neq 1$, so the theorem fails at 0. By Proposition 7.4.3 (a), it does not follow from (Assoc) and (Id) alone.

Marking scheme. (10 points) - 2: the plan (fix x_0, assume $x_0 \cdot x_0 = x_0$, take an inverse by $\exists$E). - 3: the chain of equations, with each =E naming its φ. - 2: the $\exists$E and $\forall$I side conditions checked: $\exists$E, y_0 not in $x_0 = e$, not in $\exists y\,(x_0 \cdot y = e \wedge y \cdot x_0 = e)$, in no premise and in no open assumption; $\forall$I, x_0 in no premise, in no assumption open outside its box, and not in $\forall x\,(x \cdot x = x \to x = e)$. - 1: the axioms named. - 2: the model without (Inv), with (Assoc) and (Id) checked and the theorem failing at 0. Fatal errors (0 for the derivation): an =E step that replaces a term the line does not contain; using y_0 outside its $\exists$E box.

7.4.6.3 C. Going deeper

Exercise. C1: Describe every model

Describe every model of each sentence, and prove your description.

(a) $\forall x \, \forall y \, (x = y)$.
(b) $\exists x \, \exists y \, \neg(x = y) \wedge \forall x \, \forall y \, \forall z \, (x = y \vee x = z \vee y = z)$.
(c) $\forall x \, \forall y \, R(x, y)$.
(d) $\forall x \, \exists y \, (R(x, y) \wedge \forall z \, (R(x, z) \to z = y))$.

Solution. *(a)* "Any two elements are equal": exactly the structures with one element. If D has $d \neq e$, then (S6) at d and e meets the equation $x = y$ with different values, which (S2) rejects. With one element, every equation between its values holds. Domains are non-empty, so "at most one element" is "exactly one".

(b) "There are two different elements, and among any three, two are equal": exactly the structures with two elements. The first conjunct gives two different elements $d \neq e$. If a third f differed from both, then x, y, z at d, e, f would make all three equations false. Conversely, in a two-element domain the two elements witness the first conjunct, and any three values repeat one of them.

(c) "Everything is related to everything": exactly the structures with $R^{\mathfrak{M}} = D \times D$, on any non-empty D. By (S6) twice, every pair (d, e) must be in $R^{\mathfrak{M}}$, and if every pair is, the sentence holds.

(d) "Each element is related to exactly one element": exactly the structures in which $R^{\mathfrak{M}}$ is the graph of a function from D to D, that is, each d has exactly one e with $(d, e) \in R^{\mathfrak{M}}$. For each d, (S7) gives an e with $(d, e) \in R^{\mathfrak{M}}$, and the inner $\forall z$ says that any f with $(d, f) \in R^{\mathfrak{M}}$ is e. Conversely, if each d has exactly one such e, that e is the witness, and the inner clause holds by uniqueness.

Marking scheme. (10 points) - 2.5 per part: 1 for the description, 1.5 for both directions (every model is of this kind; every structure of this kind is a model). Fatal error (0 for (a)): forgetting that domains are non-empty, so that "at most one" is "exactly one".

Exercise. C2: The partial order axioms are independent

Show that each of (Refl), (Antisym) and (Trans), for $\leq$, is independent of the other two.

Solution. For (a) of Proposition 7.4.3, three structures:

- *Not (Refl)*: $D = \{a, b\}$, $\leq^{\mathfrak{M}} = \{(a, b)\}$. Antisymmetric and transitive, vacuously: there is no pair (b, a) and no chain of two arrows. No loops.
- *Not (Antisym)*: $D = \{a, b\}$, all four pairs. Reflexive and transitive; $a \leq b$ and $b \leq a$ with $a \neq b$.
- *Not (Trans)*: $D = \{a, b, c\}$, the loops together with (a, b) and (b, c). Reflexive; antisymmetric, since no two different elements are related both ways; (a, c) is missing.

For (b), each axiom is also true in a model of the other two, for instance in $(\mathbb{N}, \leq)$, where all three hold. So each axiom is independent of the other two.

Marking scheme. (10 points) - 3 per axiom: the structure 1, the two other axioms checked 1, the failing pair 1. - 1: independence cited correctly, including that each axiom is true in some model of the other two. Fatal error (0 for the axiom): a structure in which one of the "other two" fails.

Exercise. C3: Exit check: a theory checked three ways

Let T be the theory of a strict linear order with a least element: (Irrefl), (Trans), (Tri), and

$$\text{(Least)} \quad \exists x \, \forall y \, (x = y \lor x < y).$$

(a) Show that T is consistent.

(b) Show that (Dense) is independent of T.

(c) Derive from T that the least element is unique: $\forall x \, \forall y \, ((L(x) \land L(y)) \to x = y)$, where $L(x)$ abbreviates $\forall z \, (x = z \lor x < z)$. You may use TAUT on the propositional skeleton.

Solution. *(a)* $(\mathbb{N}, <)$ is a model: no number is below itself, $<$ is transitive, any two numbers are equal or one is below the other, and 0 is a least element. By Proposition 7.4.2, T is consistent.

(b) In $(\mathbb{N}, <)$, (Dense) fails at $0 < 1$. The non-negative rationals with $<$ are a model of T (no rational is below itself, $<$ is transitive, any two rationals are equal or one is below the other, and 0 is least, since no non-negative rational is below it), in which (Dense) holds: between $p < q$ lies $(p + q)/2$. By Proposition 7.4.3, (Dense) is independent of T.

(c) Written out, $L(x_0)$ is $\forall z \, (x_0 = z \lor x_0 < z)$. The plan: fix x_0, y_0, assume both are least, instantiate each at the other, and let a TAUT step combine the two disjunctions with (Irrefl) and (Trans) at x_0, y_0.

1	$\forall x\,\neg(x < x)$				
2	$\forall x\,\forall y\,\forall z\,((x < y \wedge y < z) \to x < z)$				
3	x_0	y_0		$\forall z\,(x_0 = z \vee x_0 < z) \wedge \forall z\,(y_0 = z \vee y_0 < z)$	asm, for $\to$I
4				$\forall z\,(x_0 = z \vee x_0 < z)$	$\wedge$E 3
5				$\forall z\,(y_0 = z \vee y_0 < z)$	$\wedge$E 3
6				$x_0 = y_0 \vee x_0 < y_0$	$\forall$E 4
7				$y_0 = x_0 \vee y_0 < x_0$	$\forall$E 5
8				$\neg(x_0 < x_0)$	$\forall$E 1
9				$\forall y\,\forall z\,((x_0 < y \wedge y < z) \to x_0 < z)$	$\forall$E 2
10				$\forall z\,((x_0 < y_0 \wedge y_0 < z) \to x_0 < z)$	$\forall$E 9
11				$(x_0 < y_0 \wedge y_0 < x_0) \to x_0 < x_0$	$\forall$E 10
12				$y_0 = x_0$	asm, for $\to$I
13				$y_0 = y_0$	=I
14				$x_0 = y_0$	=E 12, 13
15				$y_0 = x_0 \to x_0 = y_0$	$\to$I 12–14
16				$x_0 = y_0$	TAUT 6, 7, 8, 11, 15
17			$(\forall z\,(x_0 = z \vee x_0 < z) \wedge \forall z\,(y_0 = z \vee y_0 < z)) \to x_0 = y_0$		$\to$I 3–16
18		$\forall y\,((\forall z\,(x_0 = z \vee x_0 < z) \wedge \forall z\,(y = z \vee y < z)) \to x_0 = y)$			$\forall$I 3–17
19	$\forall x\,\forall y\,((\forall z\,(x = z \vee x < z) \wedge \forall z\,(y = z \vee y < z)) \to x = y)$				$\forall$I 3–18

Line 14 is =E with φ the formula $z = y_0$: line 13 is $\varphi[y_0/z]$, and line 12 says $y_0 = x_0$, so the rule gives $\varphi[x_0/z]$. The TAUT step on line 16 reads the skeleton with the letters A for $x_0 = y_0$, B for $x_0 < y_0$, C for $y_0 = x_0$, D for $y_0 < x_0$ and E for $x_0 < x_0$. From $A \vee B$, $C \vee D$, $\neg E$, $(B \wedge D) \to E$ and $C \to A$, the letter A follows in every row: if A is false, then B is true; $C \to A$ makes C false, so D is true; then E is true, against $\neg E$. At line 18, y_0 is in no premise, no assumption is open outside its box, and y_0 is not in line 18; at line 19 the same holds for x_0. Only (Irrefl) and (Trans) were used.

Marking scheme. (10 points) - 2: (a), a model with every axiom checked. - 3: (b), two models of T (the axioms checked for the new one), with (Dense) checked in each. - 4: (c), the derivation: the case $y_0 = x_0$ closed, by symmetry of $=$ via =I and =E or by =E into $x_0 < y_0$ (1), the instance of (Trans) at x_0, y_0, x_0, reached by three $\forall$E steps (1), the cases combined, by a TAUT step whose skeleton is a tautology or by an equivalent $\vee$E and $\bot$E case analysis (1), the $\forall$I side conditions (1). - 1: names the axioms used. Fatal error (0 for (c)): a TAUT step that needs $y_0 = x_0$ and $x_0 = y_0$ to be the same letter (they are not).

7.5 Definability and Isomorphism

A sentence is true or false in a structure. A formula with one free variable is different: it is true of some elements of a structure and false of others, so it picks out a set. In $(\mathbb{N}; +)$ the

formula $\exists y\,(x = y + y)$ is true of $0, 2, 4, \ldots$ and false of the odd numbers, so it picks out the even numbers, although the structure has no symbol for "even". This section asks which sets a structure's language can pick out. Its first half writes formulas for sets: points, intervals, orders, the primes, a supremum. Its second half, for a second reading, asks why some sets cannot be picked out at all, and answers with a map that moves the elements of a structure without changing anything a formula can see.

Remark (Looking back). Section 4.9 asked what the connectives can say, and answered: every truth function, by a normal form. This section asks the same question of a first-order language inside one structure. There the answer was "everything"; here it depends on the structure, and a set that one structure's formulas define, another's cannot.

7.5.1 Formulas define sets

A formula defines the set of elements (or tuples of elements) that make it true.

Definition 7.5.1. Definable Set

Let $\mathfrak{M}$ be a structure with domain D, and φ a formula whose free variables are among $x_1, \ldots, x_n$, all different. The set **defined by** φ in $\mathfrak{M}$ is

$$\{(d_1, \ldots, d_n) \in D^n : \mathfrak{M}, \ell[x_1 \mapsto d_1] \cdots [x_n \mapsto d_n] \models \varphi\},$$

for any assignment ℓ. A set $S \subseteq D^n$ is **definable** in $\mathfrak{M}$ if some formula defines it.

By Lemma 7.1.1 the set does not depend on ℓ: the assignments $\ell[x_1 \mapsto d_1] \cdots [x_n \mapsto d_n]$, for different ℓ, agree on $x_1, \ldots, x_n$, and only the values of the free variables matter. For $n = 1$ the set is a set of elements, and we write it in set-builder form, $\{x \in D : \varphi\}$. Set-builder notation, used since school, is exactly this: the reading of a formula in a structure.

Two remarks on the definition. The structure is part of the question: the same formula defines different sets in different structures, and a set is definable **in** $\mathfrak{M}$, never definable outright. And the formula may use only the symbols of $\mathfrak{M}$: to define a set in $(\mathbb{N}; +)$ we may write $+$, $=$, connectives and quantifiers, and nothing else, not even 0 or $<$.

The connectives act on defined sets as the set operations do. If φ and ψ define S and T (with the same free variable x), then by (S4) and (S5) the formula $\varphi \wedge \psi$ defines $S \cap T$, $\varphi \vee \psi$ defines $S \cup T$, and $\neg\varphi$ defines $D \setminus S$, the elements not in S. So the definable sets of a structure are closed under intersection, union and complement, and a set built from definable sets by these operations is definable without further work.

Example. Sets Defined by Formulas

Each formula is read back in words.

(a) In $(\mathbb{R}; +, \cdot)$, the formula $\exists y\,(x = y \cdot y)$, "$x$ is a square", defines $[0, \infty)$: a real number is the square of a real number exactly when it is non-negative.

(b) In $(\mathbb{N}; +)$, the formula $x + x = x$, "x added to itself gives itself", defines $\{0\}$, since $d + d = d$ forces $d = 0$. The formula $\exists z\,(\neg(z + z = z) \wedge x + z = y)$, "$y$ is x plus something non-zero", defines the set of pairs (d, e) with $d < e$.

(c) In $(\mathbb{N};\cdot)$, write $U(x)$ for the formula $\forall y\,(x\cdot y = y)$, "multiplying by x changes nothing". It defines $\{1\}$: at $x = 0$ the instance $y = 1$ fails, since $0 \cdot 1 = 0$; at $x \geq 2$ the instance $y = 1$ fails too, since $x \cdot 1 = x \neq 1$; and at $x = 1$ every instance holds.

(d) In $(\mathbb{N};\cdot)$, the formula

$$\neg U(x) \wedge \forall y \forall z\,(y \cdot z = x \to (U(y) \vee U(z))),$$

"x is not 1, and whenever x is a product, one factor is 1", defines the primes. At $x = 0$ the instance $y = z = 0$ fails ($0 \cdot 0 = 0$, and neither factor is 1); at $x = 1$ the first conjunct fails; at a composite $x = yz$ with $y, z \geq 2$ the instance at that y, z fails; at a prime, every factorization has a factor 1.

Each definition builds on the last: 0 and 1 are not symbols of these structures, yet the language can name them, and once 1 is defined, "prime" can be written with it. In (d), $U(y)$ is the formula $U(x)$ with y for x and its bound y renamed, say $\forall w\,(y \cdot w = w)$, so that the substituted letter is not captured.

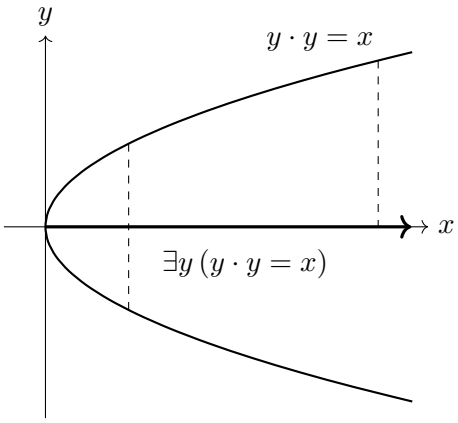
To check that a formula defines a given set, check two directions, as for any equality of sets: every element of the set satisfies the formula, and every element that satisfies the formula is in the set. In (d) the cases $x = 0$, $x = 1$ and "x composite" are the second direction, and "at a prime" is the first. Leave out a case and a definition can be wrong at one point. Delete the conjunct $\neg U(x)$ from (d): at $x = 1$ every factorization $y \cdot z = 1$ has $y = z = 1$, so the remaining formula is true of 1, and it defines the primes together with 1. Exactly one element changed sides.

Quick check. What does $\exists y\,(x = y + y)$ define in $(\mathbb{Z};+)$? In $(\mathbb{Q};+)$?

Solution. In $(\mathbb{Z};+)$, the even integers. In $(\mathbb{Q};+)$, all of $\mathbb{Q}$: every rational x is $y + y$ for $y = x/2$, which is again rational. Same formula, two structures, two sets.

7.5.2 $\exists$ is a projection

A formula with two free variables defines a set of pairs, and a set of pairs of reals can be drawn. In $(\mathbb{R};\cdot)$, the formula $y \cdot y = x$ defines $\{(d, e) : e^2 = d\}$, a parabola lying on its side. The formula $\exists y\,(y \cdot y = x)$ has the one free variable x, and by (S7) it is true of d exactly when some e puts the point (d, e) on the parabola. So it defines the set of first coordinates of the points of the parabola: its **shadow** on the x-axis, $[0, \infty)$.



The dashed segments drop two points, $(1, 1)$ and $(1, -1)$, and the point $(4, 2)$, onto their shadows 1 and 4. No point of the parabola lies above a negative x, so no negative number is in the shadow.

In general, $\exists y\, \varphi$ defines the **projection** of the set that φ defines: d is in it exactly when some e puts (d, e) in φ's set, by (S7). Dually, $\forall y\, \varphi$ keeps the d whose whole vertical line lies in φ's set, by (S6). This is the picture behind every $\exists$ in a definition: to show a set definable, it is often easiest to describe a set of pairs and project.

The two readings are tied by negation. By (S4), (S6) and (S7), $\forall y\, \varphi$ is satisfied exactly when $\exists y\, \neg\varphi$ is not, so the set defined by $\forall y\, \varphi$ is the complement of the projection of the complement of φ's set: remove φ's set from the plane, take the shadow of what is left, and keep the points of the axis that are not in that shadow.

Quick check. In $(\mathbb{R}; \leq, \cdot)$, which set does $\forall y\, (x \leq y \cdot y)$ define?

Solution. $(-\infty, 0]$. The least value of $y \cdot y$ is 0, at $y = 0$, so $x \leq y^2$ holds for every y exactly when $x \leq 0$. In the picture: the vertical line at x lies wholly in the region $x \leq y^2$ when $x \leq 0$, and the line at a positive x leaves it near $y = 0$.

7.5.3 Define it by a formula

A common task in mathematics is to turn a notion described in words into a formula, and the skill is the reverse of reading back. The notion must be said using only the symbols of the structure, and every variable that is not the defined set's own must be bound.

Example. Maximum, Maximal, Supremum

(a) In a structure with one two-place symbol $\leq$ read as a partial order: "x is a **greatest** element" is $\forall y\, (y \leq x)$, "everything is below x". "x is a **maximal** element" is $\forall y\, (x \leq y \to y = x)$, "nothing is above x except x itself". In the divisibility order on $\{1, 2, 3\}$ ($d \leq e$ when d divides e), 2 and 3 are maximal, and there is no greatest element, since 2 does not divide 3 nor 3 divide 2. On $\{1, 2, 3, 6\}$, 6 is both maximal and greatest.

(b) Add a one-place symbol A for a set of reals, and work in $(\mathbb{R}; \leq, A)$. "x is an upper bound of A" is the formula $B(x)$: $\forall y\, (A(y) \to y \leq x)$, "every element of A is at most x". "x is the supremum of A" is $B(x) \wedge \forall z\, (B(z) \to x \leq z)$, "$x$ is an upper bound, and below every upper bound". For $A = [0, 1)$, the supremum is 1: it is an upper bound,

and any $z < 1$ is exceeded by an element of $[0, 1)$, such as the larger of 0 and $(z+1)/2$, so z is not one. No element of A is greatest.

(c) "A has a greatest element" is the sentence $\exists x\,(A(x) \wedge B(x))$, "some element of A is an upper bound of A". It is false for $A = [0, 1)$ and true for $A = [0, 1]$.

In (c) the formula $B(x)$ is used inside a larger formula, as a word is used in a sentence. This is how definitions accumulate: each defined notion is an abbreviation, and the full formula is the abbreviation written out.

Defining a notion goes in three steps. First, say the notion in words that use only the structure's symbols: "upper bound" becomes "every element of A is at most x". Second, write each auxiliary notion as a formula with its own free variable, and substitute it where it is used, renaming bound variables so that nothing is captured. Third, check both directions on a small case. The commonest slip is a variable left free. Written without its $\forall y$, the formula $A(y) \to y \leq x$ has two free variables and defines a set of **pairs**, those (d, e) with $d \notin A$ or $d \leq e$; it is not a property of x at all.

A defined notion can also be given its own symbol, as "even" and "the supremum of A" are in ordinary mathematics. Section 5.8 showed (Proposition 5.8.2) that such a definition is safe: introducing a predicate symbol by $\forall x\,(P(x) \leftrightarrow \varphi)$, or a constant for an element that φ is proved to have, proves nothing new about the old symbols. Definability says which notions can be so introduced in a given structure: exactly those some formula of its language defines.

Warning. Greatest is not maximal. A maximal element has nothing strictly above it; a greatest element is above everything. In the divisibility order on $\{1, 2, 3\}$, 2 is maximal and not greatest. Writing $\forall y\,(y \leq x)$ for "maximal" defines the wrong set: in that order it defines $\emptyset$, while the maximal elements are 2 and 3.

Quick check. In the divisibility order on $\{1, 2, 3, 4, 6\}$, which elements are maximal? Is there a greatest element?

Solution. 4 and 6: neither divides any other element of the set. There is no greatest element: 4 does not divide 6, and 6 does not divide 4.

7.5.4 Isomorphic structures (second pass)

This half of the section is for a second reading. It asks when two structures are the same except for the names of their elements, and proves that such structures satisfy the same sentences. From that comes a way to show that a set is **not** definable. It uses functions and bijections in their everyday sense: a rule that sends each element to one element, and, for a bijection, one that matches the two domains one to one, every element of the second domain hit exactly once.

Definition 7.5.2. Isomorphism of Structures

Let $\mathfrak{M}$ and $\mathfrak{N}$ be structures for the same symbols, with domains D and E. An **isomorphism** $h\colon \mathfrak{M} \to \mathfrak{N}$ is a bijection $h\colon D \to E$ such that

(i) $h(c^{\mathfrak{M}}) = c^{\mathfrak{N}}$ for each constant c;

(ii) $h(f^{\mathfrak{M}}(d_1, \ldots, d_n)) = f^{\mathfrak{N}}(h(d_1), \ldots, h(d_n))$ for each function symbol f of arity n and all $d_1, \ldots, d_n \in D$;

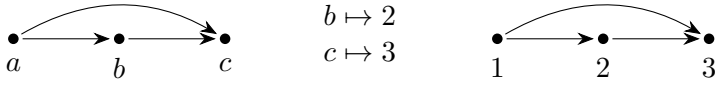
(iii) $(d_1, \ldots, d_n) \in R^{\mathfrak{M}}$ if and only if $\big(h(d_1), \ldots, h(d_n)\big) \in R^{\mathfrak{N}}$, for each predicate symbol R of arity n and all $d_1, \ldots, d_n \in D$.

An isomorphism $\mathfrak{M} \to \mathfrak{M}$ is an **automorphism** of $\mathfrak{M}$. $\mathfrak{M}$ and $\mathfrak{N}$ are **isomorphic** if an isomorphism $\mathfrak{M} \to \mathfrak{N}$ exists.

In words: h renames every element, and the renaming carries each constant to the corresponding constant, each operation to the corresponding operation, and each relation onto the corresponding relation, in both directions.

Clause (iii) is needed in both directions. The identity map $1 \mapsto 1$, $2 \mapsto 2$, from $\{1, 2\}$ with R empty to $\{1, 2\}$ with $R = \{(1, 2)\}$, is a bijection, and it sends every pair in the first relation into the second, since there are none. Yet the sentence $\exists x\, \exists y\, R(x, y)$ is false in the first structure and true in the second. The converse half of (iii) fails at $(1, 2)$, and with it the map is not an isomorphism.

Examples. $(\mathbb{N}, <)$ and the even numbers with $<$ are isomorphic, by $h(d) = 2d$: every even number is $2d$ for exactly one d, and $d < e$ if and only if $2d < 2e$. A finite example, with the bijection as a table:



On the left, $\mathfrak{M}$ has $D = \{a, b, c\}$ and $R^{\mathfrak{M}} = \{(a, b), (b, c), (a, c)\}$; on the right is $(\{1, 2, 3\}, <)$, with an arrow from d to e when $d < e$. The table is a bijection, and it sends the three arrows on the left exactly onto the three on the right. A non-example: $d \mapsto d + 1$, from $(\mathbb{N}, <)$ to itself, preserves $<$ in both directions and is one to one, but it is not onto: nothing is sent to 0. So it is not an automorphism.

Theorem 7.5.1. Isomorphic Structures Satisfy the Same Sentences

Let $h\colon \mathfrak{M} \to \mathfrak{N}$ be an isomorphism. For every formula φ and every assignment ℓ in $\mathfrak{M}$,

$$\mathfrak{M}, \ell \models \varphi \quad \text{if and only if} \quad \mathfrak{N}, h \circ \ell \models \varphi,$$

where $h \circ \ell$ gives each variable x the value $h(\ell(x))$. In particular, $\mathfrak{M} \models \sigma$ if and only if $\mathfrak{N} \models \sigma$, for every sentence σ.

In words: renaming the elements, consistently, changes the truth of nothing.

The theorem is most often used backwards: a sentence true in one structure and false in another shows that the two are **not** isomorphic. The sentence $\forall x\, \exists y\, (y + y = x)$, "every element is a sum of two equal ones", is true in $(\mathbb{Q}; +)$ (take $y = x/2$) and false in $(\mathbb{Z}; +)$ (at $x = 1$), so no bijection $\mathbb{Q} \to \mathbb{Z}$ carries addition to addition, although bijections $\mathbb{Q} \to \mathbb{Z}$ exist (*Sets, Relations, and Cardinality*, Chapter 3). The converse of the theorem fails: two structures can satisfy the same sentences without being isomorphic, and the optional Section 7.6 meets such a pair.

The proof is in the fold below.

↻ **Second reading** · *The proof: terms, then formulas*

Idea. An induction on terms and then on formulas, of the kind Section 7.1 gave for Lemma 7.1.1. Each clause of the definition of isomorphism carries one kind of atomic formula across; injectivity carries equations back; and in the quantifier cases, as d runs over D, its image $h(d)$ must run over all of E, which is what "onto" gives.

Proof. Terms. First, $h(t^{\mathfrak{M},\ell}) = t^{\mathfrak{N},h\circ\ell}$ for every term t, by induction on t. For a variable x, both sides are $h(\ell(x))$, by the definition of $h \circ \ell$. For a constant, this is clause (i). For $t = f(t_1, \dots, t_n)$, clause (ii) and the hypothesis for $t_1, \dots, t_n$ give

$$h(f^{\mathfrak{M}}(t_1^{\mathfrak{M},\ell}, \dots, t_n^{\mathfrak{M},\ell})) = f^{\mathfrak{N}}(h(t_1^{\mathfrak{M},\ell}), \dots, h(t_n^{\mathfrak{M},\ell})) = f^{\mathfrak{N}}(t_1^{\mathfrak{N},h\circ\ell}, \dots, t_n^{\mathfrak{N},h\circ\ell}).$$

Formulas. Structural induction on φ, with the claim stated for **all** assignments ℓ at once, since the quantifier clauses change the assignment.

Atomic. By (S1), $\mathfrak{M}, \ell \models R(t_1, \dots, t_n)$ means that the tuple of the values $t_i^{\mathfrak{M},\ell}$ is in $R^{\mathfrak{M}}$. By clause (iii) that holds exactly when the tuple of their images is in $R^{\mathfrak{N}}$, and by the claim for terms the images are the values $t_i^{\mathfrak{N},h\circ\ell}$: this is (S1) for $\mathfrak{N}, h \circ \ell$. *Equations.* $t_1^{\mathfrak{M},\ell} = t_2^{\mathfrak{M},\ell}$ implies $h(t_1^{\mathfrak{M},\ell}) = h(t_2^{\mathfrak{M},\ell})$; the converse holds because h is **injective**. With the claim for terms, this is (S2) on both sides. $\bot$ is satisfied on neither side, by (S3).

Connectives. The answers for φ and ψ agree on the two sides by hypothesis, and (S4), (S5) combine them by the same rule on both sides.

Universal quantifier. $\mathfrak{M}, \ell \models \forall x\, \varphi$ holds if and only if $\mathfrak{M}, \ell[x \mapsto d] \models \varphi$ for every $d \in D$, by (S6); by the hypothesis, applied to the assignment $\ell[x \mapsto d]$, if and only if $\mathfrak{N}, h\circ(\ell[x \mapsto d]) \models \varphi$ for every $d \in D$. Now $h \circ (\ell[x \mapsto d]) = (h \circ \ell)[x \mapsto h(d)]$: both give x the value $h(d)$ and every other variable y the value $h(\ell(y))$. As d runs over D, $h(d)$ runs over **all** of E, because h is **onto**. So the condition says that $\mathfrak{N}, (h \circ \ell)[x \mapsto e] \models \varphi$ for every $e \in E$, which is (S6) for $\mathfrak{N}, h \circ \ell \models \forall x\, \varphi$.

Existential quantifier. The same, with (S7): some $d \in D$ works on the left exactly when some $e = h(d) \in E$ works on the right, and every $e \in E$ is such an $h(d)$, since h is **onto**.

For a sentence, Lemma 7.1.1 makes the assignment irrelevant on both sides, which gives the last sentence of the statement.

■

Quick check. Why can no sentence in the symbol $<$ tell $(\mathbb{N}, <)$ from the even numbers with $<$?

Solution. Doubling is an isomorphism between them, so by Theorem 7.5.1 they satisfy the same sentences.

7.5.5 What an automorphism cannot move (second pass)

An automorphism is a renaming of a structure onto itself. By the theorem, a formula cannot tell an element from its image.

Corollary 7.5.2. Automorphisms Preserve Definable Sets

If h is an automorphism of $\mathfrak{M}$ and $S \subseteq D$ is definable in $\mathfrak{M}$, then for every $d \in D$, $d \in S$ if and only if $h(d) \in S$.

Proof. Let φ, with free variable x, define S, and let ℓ be an assignment with $\ell(x) = d$. By Theorem 7.5.1 with $\mathfrak{N} = \mathfrak{M}$, $\mathfrak{M}, \ell \models \varphi$ if and only if $\mathfrak{M}, h \circ \ell \models \varphi$. The first says $d \in S$; since $(h \circ \ell)(x) = h(d)$, the second says $h(d) \in S$. ■

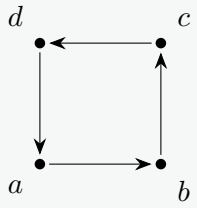
So to show that a set is **not** definable, find an automorphism that moves an element of it to an element outside it.

Example. Zero Cannot Be Defined from the Order of the Integers

In $(\mathbb{Z}, <)$, the shift $h(d) = d+1$ is an automorphism: it is a bijection, with inverse $d \mapsto d-1$, and $d < e$ if and only if $d+1 < e+1$. It sends 0 to 1, and $1 \notin \{0\}$. By Corollary 7.5.2, $\{0\}$ is not definable in $(\mathbb{Z}, <)$: no formula in $<$ alone singles out zero among the integers.

Contrast $(\mathbb{N}, <)$. There $\{0\}$ is definable, by $\forall y \neg(y < x)$, "nothing is below x". The corollary is not contradicted, since $d \mapsto d+1$ is not an automorphism of $(\mathbb{N}, <)$: it is not onto.

A finite picture: the four-cycle below, with $D = \{a, b, c, d\}$ and R the four arrows $a \to b \to c \to d \to a$.



The rotation $a \mapsto b$, $b \mapsto c$, $c \mapsto d$, $d \mapsto a$ is a bijection that sends each arrow to an arrow and each non-arrow to a non-arrow, so it is an automorphism. Applied once, twice or three times, it moves any element to any other, so no single element is definable.

A minimal change makes one element definable again. Add the loop (a, a) to R. Now $R(x, x)$ defines $\{a\}$, and the rotation is no longer an automorphism, since it sends the arrow (a, a) to (b, b), which is not an arrow. The method can only use the automorphisms a structure has;

each of the three non-trivial rotations sends the loop (a, a) to a loop at b, c or d, so adding one arrow removed all three.

Warning. An automorphism shows that a set is **not** definable; it never shows that one is. That no automorphism moves a set is not evidence that a formula defines it. To show a set definable, write the formula.

Quick check. In the four-cycle, where does the rotation send the set $\{a, c\}$? Could $\{a, c\}$ be definable?

Solution. The rotation sends a to b and c to d, so it moves $a \in \{a, c\}$ to $b \notin \{a, c\}$. By Corollary 7.5.2, $\{a, c\}$ is not definable. Its complement $\{b, d\}$ is not definable either, by the same rotation.

Remark (Looking ahead). *Number Systems: From $\mathbb{N}$ to $\mathbb{R}$*, Chapter 6, proves that any two complete ordered fields are isomorphic, by an isomorphism built from the rationals below each element; by this section, they then satisfy the same sentences.

7.5.6 Exercises

7.5.6.1 A. Check your understanding

Exercise. A1

(a) Define "$S \subseteq D$ is definable in $\mathfrak{M}$".

(b) Write formulas defining, in $(\mathbb{N}; +)$, the even numbers and the set $\{0\}$.

(c) Why does a formula with two free variables define a set of pairs, and not a set of elements?

(d) *(Second pass.)* State Theorem 7.5.1.

(e) *(Second pass; key proof.)* Where does the proof of Theorem 7.5.1 use that h is onto, and where that it is injective? What goes wrong for $d \mapsto d + 1$ on $(\mathbb{N}, <)$?

(f) *(Second pass.)* Why does $x \mapsto x + 1$ show $\{0\}$ undefinable in $(\mathbb{Z}, <)$ but not in $(\mathbb{N}, <)$?

Solution. *(a)* Some formula φ with free variable x (at most) defines it: $S = \{d \in D : \mathfrak{M}, \ell[x \mapsto d] \models \varphi\}$, for any assignment ℓ (Definition 7.5.1).

(b) $\exists y\,(x = y + y)$, "x is a sum of two equal numbers"; $x + x = x$, "x added to itself is itself".

(c) Its truth depends on the values of both variables (and, by Lemma 7.1.1, on nothing else), so it is true or false of pairs of values.

(d) If $h \colon \mathfrak{M} \to \mathfrak{N}$ is an isomorphism, then $\mathfrak{M}, \ell \models \varphi$ if and only if $\mathfrak{N}, h \circ \ell \models \varphi$, for every formula φ and every assignment ℓ in $\mathfrak{M}$; in particular $\mathfrak{M}$ and $\mathfrak{N}$ satisfy the same sentences.

(e) Onto: in the $\forall$ and $\exists$ cases, so that $h(d)$ runs over all of E as d runs over D. Injective: in the equation case, to go from $h(t_1^{\mathfrak{M},\ell}) = h(t_2^{\mathfrak{M},\ell})$ back to $t_1^{\mathfrak{M},\ell} = t_2^{\mathfrak{M},\ell}$. The map $d \mapsto d+1$ is not onto, and the $\forall$ case fails: $\forall y\, \neg(y < x)$ is satisfied when x has the value 0, and not when x has the value $h(0) = 1$, because the element 0 below 1 is exactly the one the map never reaches.

(f) It is an automorphism of $(\mathbb{Z}, <)$ and moves 0; it is not an automorphism of $(\mathbb{N}, <)$ (not onto), and in $(\mathbb{N}, <)$ the set $\{0\}$ is in fact definable, by $\forall y\, \neg(y < x)$.

Exercise. A2: Drill

Each part takes a line or two. Parts (a)–(f) belong to the first half of the section; parts (g)–(l) to the second half (second pass). They run from reading a formula, through writing one, to judging maps and using them.

(a) What subset of $\mathbb{Z}$ does $\exists y\, (x = y + y + y)$ define in $(\mathbb{Z}; +)$?
(b) Define $\{0\}$ in $(\mathbb{Z}; +)$.
(c) Define the pairs with $x \leq y$ in $(\mathbb{N}; +)$.
(d) In $(\{1, 2, 3\}, <)$, which set does $\forall y\, \neg(y < x)$ define? In $(\mathbb{Z}, <)$?
(e) Write "x is maximal" and "x is greatest" for a partial order $\leq$.
(f) In $(\mathbb{R}; 1, +, \cdot)$, φ is $x \cdot x + y \cdot y = 1$. What does $\exists y\, \varphi$ define?
(g) Is $h(d) = 2d$ an isomorphism from $(\mathbb{Z}, <)$ to the even integers with $<$?
(h) Is $h(d) = -d$ an automorphism of $(\mathbb{Z}, <)$? Of $(\mathbb{Z}; +)$?
(i) Using (h), is $\{1\}$ definable in $(\mathbb{Z}; +)$?
(j) Is that consistent with (b)?
(k) $\mathfrak{M}$ has $D = \{a, b, c\}$ and $R^{\mathfrak{M}} = \{(a, c), (b, c)\}$. Which singletons are definable?
(l) Are $(\mathbb{N}, <)$ and $(\mathbb{Z}, <)$ isomorphic? Give a sentence that shows it.

Solution.

(a) The multiples of 3.
(b) $x + x = x$.
(c) $\exists z\, (x + z = y)$, "y is x plus something".
(d) $\{1\}$, "nothing is below x"; $\emptyset$, since every integer has a smaller one.
(e) $\forall y\, (x \leq y \to y = x)$; $\forall y\, (y \leq x)$.
(f) $[-1, 1]$, the shadow of the unit circle on the x-axis.
(g) Yes: a bijection onto the evens, and $d < e$ if and only if $2d < 2e$.
(h) Not of $(\mathbb{Z}, <)$: $0 < 1$, but $-0 = 0 > -1$. Yes of $(\mathbb{Z}; +)$: it is a bijection (its own inverse), and $-(d + e) = (-d) + (-e)$.
(i) No: the automorphism $d \mapsto -d$ sends 1 to $-1 \notin \{1\}$.
(j) Yes: $d \mapsto -d$ sends 0 to 0, so it does not move $\{0\}$.
(k) $\{c\}$, by $\exists y\, R(y, x)$, "something points to x". Not $\{a\}$ or $\{b\}$: swapping a and b (and fixing c) is an automorphism.
(l) No. $\exists x\, \forall y\, \neg(y < x)$, "there is a least element", is true in $(\mathbb{N}, <)$ and false in $(\mathbb{Z}, <)$;

isomorphic structures would agree on it.

7.5.6.2 B. Practice

Exercise. B1: Define it by a formula

Define by a formula, and read each formula back in words. Parts (i)–(iii) are in $(\mathbb{N}; \cdot)$; parts (iv) and (v) are in $(\mathbb{R}; \leq, A)$, with A a one-place symbol.

(i) "x divides y".

(ii) "x is a square".

(iii) "x and y have no common divisor except 1".

(iv) "A is bounded above".

(v) "A has a supremum that is not in A".

Solution.

(i) $\exists z\,(x \cdot z = y)$: "$y$ is x times something".

(ii) $\exists y\,(y \cdot y = x)$: "$x$ is something times itself".

(iii) $\forall z\,((\exists u\,(z \cdot u = x) \wedge \exists v\,(z \cdot v = y)) \to \forall w\,(z \cdot w = w))$: "every common divisor z of x and y is 1". The consequent is $U(z)$ of Example (Sets Defined by Formulas) (c), with its bound variable renamed to w: written with the bound y of U, it would capture the free y. (At $x = y = 0$ every z is a common divisor, so the formula fails, as it should.)

(iv) $\exists x\,B(x)$, with $B(x)$ the upper-bound formula $\forall y\,(A(y) \to y \leq x)$: "some x is at least every element of A".

(v) $\exists x\,(B(x) \wedge \forall z\,(B(z) \to x \leq z) \wedge \neg A(x))$: "some least upper bound of A is not in A".

Marking scheme. (10 points) 2 for each part: 1 for the formula, 1 for its reading back, with any auxiliary formula (U, B) written out or cited. Fatal error (0 for the part): a free variable left over that is not one of the defined set's.

Exercise. B2: What does this formula define?

Describe the set each formula defines.

(i) $\exists y\,(x = y \cdot y \cdot y)$ in $(\mathbb{R}; \cdot)$.

(ii) The same formula in $(\mathbb{Z}; \cdot)$.

(iii) $\forall y\,(y \cdot x = x)$ in $(\mathbb{N}; \cdot)$.

(iv) $\exists z\,(\neg(z + z = z) \wedge x + z = y)$ in $(\mathbb{Z}; +)$.

(v) $\exists y\,(x < y \wedge \forall z\,\neg(x < z \wedge z < y))$ in $(\mathbb{Q}, <)$, and in $(\mathbb{Z}, <)$.

Solution.

(i) All of $\mathbb{R}$: every real number has a real cube root.

(ii) The cubes, $\{\ldots, -8, -1, 0, 1, 8, 27, \ldots\}$.

(iii) $\{0\}$: at $x = 0$, $y \cdot 0 = 0$ for every y; at $x \geq 1$, the instance $y = 0$ fails.

(iv) All pairs (x, y) with $x \neq y$. In $\mathbb{Z}$, $z = y - x$ always exists, and it is non-zero exactly when $x \neq y$. The same formula defined $x < y$ in $(\mathbb{N}; +)$ (Example (Sets Defined by Formulas) (b)); in $\mathbb{Z}$, where a negative z is available, it defines something else.

(v) In $(\mathbb{Q}, <)$, $\emptyset$: between x and any larger y lies $(x + y)/2$. In $(\mathbb{Z}, <)$, all of $\mathbb{Z}$: take $y = x + 1$.

Marking scheme. (10 points) 2 for each part; in (iv), full marks require noticing that $\mathbb{Z}$ is not $\mathbb{N}$. Fatal error (0 for the part): an answer that ignores the structure, such as the same set for both structures in (v).

Exercise. B3: Automorphisms of small structures

(Second pass.)

(a) Find every automorphism of $\mathfrak{M}$: $D = \{1, 2, 3, 4\}$, $R^{\mathfrak{M}} = \{(1, 1), (1, 2), (2, 1), (2, 2), (3, 3)\}$, the answer given to Exercise (B1: Construct a relation, or show none exists, used later) (ii).

(b) Which singletons and which two-element subsets of D does that leave possibly definable? Define those that are.

(c) Is $d \mapsto d + 1$ an automorphism of $(\mathbb{Z}; <, 0)$, with 0 a constant?

Solution. *(a)* An automorphism sends related pairs to related pairs and unrelated pairs to unrelated pairs. 4 is the only element related to nothing, so it must go to 4. 3 is the only element related to itself and to nothing else, so it must go to 3. That leaves 1 and 2, which may stay or be swapped, and both choices preserve R. So there are two automorphisms: the identity, and the swap of 1 and 2.

(b) $\{1\}$ and $\{2\}$ are not definable: the swap moves them. $\{3\}$ is defined by $R(x, x) \wedge \forall y\,(R(x, y) \to y = x)$; $\{4\}$ by $\neg R(x, x)$. Among the two-element sets, $\{1, 2\}$ is defined by $\exists y\,(R(x, y) \wedge \neg(y = x))$, "$x$ is related to something else", and $\{3, 4\}$ by its negation $\neg\exists y\,(R(x, y) \wedge \neg(y = x))$. The other four, $\{1, 3\}$, $\{1, 4\}$, $\{2, 3\}$, $\{2, 4\}$, are moved by the swap, so none is definable.

(c) No: it sends the value 0 of the constant to 1, and clause (i) of Definition 7.5.2 asks $h(0) = 0$.

Marking scheme. (10 points) 4 for (a): both automorphisms, with the argument that 3 and 4 are fixed. 4 for (b): each definable set with its formula, and each other set with the automorphism that moves it. 2 for (c). Fatal error (0 for (b)): claiming a set definable because no automorphism moving it was found.

7.5.6.3 C. Going deeper

Exercise. C1: The definable subsets of $(\mathbb{Z}, <)$

(Second pass.)

(a) Prove that the only subsets of $\mathbb{Z}$ definable in $(\mathbb{Z}, <)$ by a formula with one free variable are $\emptyset$ and $\mathbb{Z}$.

(b) Show that the relation "$y = x + 1$" is definable in $(\mathbb{Z}, <)$.

(c) Why does (a) not contradict (b)?

Solution. *(a)* $\emptyset$ and $\mathbb{Z}$ are defined by $\neg(x = x)$ and $x = x$. Conversely, let S be definable and non-empty, with $n \in S$, and let $m \in \mathbb{Z}$. The shift $d \mapsto d + (m - n)$ is an automorphism of $(\mathbb{Z}, <)$: it is a bijection, with inverse $d \mapsto d - (m - n)$, and $d < e$ if and only if $d + (m - n) < e + (m - n)$. It sends n to m, so $m \in S$ by Corollary 7.5.2. Since m was arbitrary, $S = \mathbb{Z}$.

(b) $x < y \wedge \neg\exists z\,(x < z \wedge z < y)$, "$y$ is above x with nothing in between". In $\mathbb{Z}$ it holds of (d, e) exactly when $d < e$ and no integer lies strictly between, that is, $e = d + 1$: if $e = d + 1$, an integer z with $d < z < d + 1$ does not exist; if $e \geq d + 2$, then $z = d + 1$ lies between.

(c) The corollary, and the argument of (a), concern sets of elements. "$y = x + 1$" defines a set of pairs, and a shift moves each pair $(n, n + 1)$ to another pair $(m, m + 1)$ of the same set, so the set of pairs is not moved. What (a) forbids is a formula singling out some integers and not others; (b) singles out a relation that every integer has to its successor.

Marking scheme. (10 points) 5 for (a): the shift shown to be an automorphism (bijective and order-preserving), then the corollary applied. 3 for (b), with both directions checked in $\mathbb{Z}$. 2 for (c). Fatal error (0 for (a)): using $d \mapsto -d$, which does not preserve $<$.

Exercise. C2: Challenge: the automorphisms of $(\mathbb{Q}, <)$

(Second pass.)

(a) Show that $x \mapsto 2x$, and $x \mapsto x + q$ for each rational q, are automorphisms of $(\mathbb{Q}, <)$.

(b) Deduce that the only subsets of $\mathbb{Q}$ definable in $(\mathbb{Q}, <)$ by a formula with one free variable are $\emptyset$ and $\mathbb{Q}$.

(c) Show that $\{x \in \mathbb{Q} : x > 0\}$ is definable in $(\mathbb{Q}; <, 0)$, with 0 a constant, but not in $(\mathbb{Q}, <)$.

Solution. *(a)* Both are bijections of $\mathbb{Q}$, with inverses $x \mapsto x/2$ and $x \mapsto x - q$, and both preserve $<$ in both directions: $x < y$ if and only if $2x < 2y$, if and only if $x + q < y + q$.

(b) The translations alone suffice: given p in a non-empty definable set S and any $p' \in \mathbb{Q}$, the automorphism $x \mapsto x + (p' - p)$ sends p to p', so $p' \in S$ by Corollary 7.5.2. As in C1, $S = \mathbb{Q}$.

(c) In $(\mathbb{Q}; <, 0)$ the formula $0 < x$ defines it. In $(\mathbb{Q}, <)$ it is neither $\emptyset$ nor $\mathbb{Q}$, so by (b) it is not definable.

Marking scheme. (10 points) 3 for (a), 4 for (b), 3 for (c). Fatal error (0 for (b)): an argument with $x \mapsto 2x$ alone, which fixes 0 and so cannot move 0 to anything else.

7.5.6.4 R. Review

Exercise. R1

Decide whether each sequent is derivable, and justify your answer.

(a) $(p \to q) \to r \vdash p \to r$.
(b) $(p \to q) \to r \vdash q \to r$.

Solution. (a) Not derivable. Look for a refuting row, backwards from the conclusion: $p \to r$ is false only when $p = \mathrm{T}$ and $r = \mathrm{F}$. The premise must then be true, so its hypothesis $p \to q$ must be false, which needs $q = \mathrm{F}$. In the row $p = \mathrm{T}$, $q = \mathrm{F}$, $r = \mathrm{F}$, the premise is $\mathrm{F} \to \mathrm{F}$, true, and the conclusion is $\mathrm{T} \to \mathrm{F}$, false. By Corollary 4.7.3 there is no derivation.

(b) Derivable. The goal is an implication, so open a box on q; the premise is waiting for $p \to q$, which a box on p produces by repeating q (the goal-directed and premise-directed steps of the search method of Section 3.1).

1	$(p \to q) \to r$	
2	$\quad q$	asm, for $\to$I
3	$\quad\quad p$	asm, for $\to$I
4	$\quad\quad q$	R 2
5	$\quad p \to q$	$\to$I 3–4
6	$\quad r$	$\to$E 1, 5
7	$q \to r$	$\to$I 2–6

The row method of (a) could not settle (b): no row refutes it, and only a derivation (or Theorem 4.8.4) shows it is derivable.

Exercise. R2

Derive $\neg(P \to \neg Q) \vdash \neg\neg P \wedge \neg\neg Q$ without RAA and without derived rules that use it. Which rule would you need to derive $P \wedge Q$ itself?

Solution. The goal is a conjunction, so derive each conjunct. Each is a negation, so each gets a $\neg$I box, whose goal is $\bot$. The only way to use the premise $\neg(P \to \neg Q)$ is to clash with it, so inside each box we build $P \to \neg Q$. In the first box, from $\neg P$: a $\to$I box on P clashes P with $\neg P$ and uses $\bot$E (the pattern of Proposition 2.1.2). In the second, from $\neg Q$: a $\to$I box on P simply repeats $\neg Q$. These are the rules of Chapters 1 and 2 (Definition 2.2.1, Definition 2.1.1, Definition 1.4.1).

1	$\neg(P \to \neg Q)$	
2	$\quad \neg P$	asm, for $\neg$I
3	$\quad\quad P$	asm, for $\to$I
4	$\quad\quad \bot$	$\neg$E 3, 2
5	$\quad\quad \neg Q$	$\bot$E 4
6	$\quad P \to \neg Q$	$\to$I 3–5
7	$\quad \bot$	$\neg$E 6, 1
8	$\neg\neg P$	$\neg$I 2–7
9	$\quad \neg Q$	asm, for $\neg$I
10	$\quad\quad P$	asm, for $\to$I
11	$\quad\quad \neg Q$	R 9
12	$\quad P \to \neg Q$	$\to$I 10–11
13	$\quad \bot$	$\neg$E 12, 1
14	$\neg\neg Q$	$\neg$I 9–13
15	$\neg\neg P \wedge \neg\neg Q$	$\wedge$I 8, 14

No RAA is used. To reach $P \wedge Q$ we would need P from $\neg\neg P$ and Q from $\neg\neg Q$, that is, $\neg\neg$E, which is derived with RAA (Proposition 2.5.1), or an RAA box directly (Definition 2.5.1). The look-alike is Exercise (B1: Three searches) (b), $\neg(P \to Q) \vdash P \wedge \neg Q$: there the conjunct P needed RAA, while $\neg Q$ came by $\neg$I alone. Here both conjuncts stop at a double negation. That no derivation of P (or of $P \wedge Q$) from the premise avoids RAA is true, but we state it without proof, and nothing depends on it.

Exercise. R3

Derive $\neg\exists x\,(P(x) \to \neg Q(x)) \vdash \forall x\,(\neg\neg P(x) \wedge \neg\neg Q(x))$ without RAA.

Solution. The goal is $\forall x$, so open a flagged box with a fresh name x_0 (Definition 5.4.1); the new goal is $\neg\neg P(x_0) \wedge \neg\neg Q(x_0)$. Inside, run the derivation of R2 (Exercise (R2)) with $P(x_0)$ and $Q(x_0)$ in place of P and Q. One change: the premise is now $\neg\exists x\,(P(x) \to \neg Q(x))$, not $\neg(P \to \neg Q)$, so each clash first turns $P(x_0) \to \neg Q(x_0)$ into $\exists x\,(P(x) \to \neg Q(x))$ by $\exists$I at x_0 (Definition 5.5.1).

1		$\neg\exists x\,(P(x) \to \neg Q(x))$	
2	x_0	$\neg P(x_0)$	asm, for $\neg$I
3		$P(x_0)$	asm, for $\to$I
4		$\bot$	$\neg$E 3, 2
5		$\neg Q(x_0)$	$\bot$E 4
6		$P(x_0) \to \neg Q(x_0)$	$\to$I 3–5
7		$\exists x\,(P(x) \to \neg Q(x))$	$\exists$I 6
8		$\bot$	$\neg$E 7, 1
9		$\neg\neg P(x_0)$	$\neg$I 2–8
10		$\neg Q(x_0)$	asm, for $\neg$I
11		$P(x_0)$	asm, for $\to$I
12		$\neg Q(x_0)$	R 10
13		$P(x_0) \to \neg Q(x_0)$	$\to$I 11–12
14		$\exists x\,(P(x) \to \neg Q(x))$	$\exists$I 13
15		$\bot$	$\neg$E 14, 1
16		$\neg\neg Q(x_0)$	$\neg$I 10–15
17		$\neg\neg P(x_0) \wedge \neg\neg Q(x_0)$	$\wedge$I 9, 16
18		$\forall x\,(\neg\neg P(x) \wedge \neg\neg Q(x))$	$\forall$I 2–17

The side condition of $\forall$I at line 18: x_0 does not appear in the premise on line 1, no box is open at line 18, and x_0 does not appear in the conclusion. The flagged box immediately contains the $\neg$I box of line 2, so it is cited from line 2, the first line of that inner box. Lines 7 and 14 are legal $\exists$I steps: $P(x_0) \to \neg Q(x_0)$ is $(P(x) \to \neg Q(x))[x_0/x]$. The whole derivation is R2's, run inside a flagged box, with an $\exists$I before each clash.

Exercise. R4

Is the sentence $\forall x\,\exists y\,(x < y) \to \exists y\,\forall x\,(x < y)$ true in $(\mathbb{N}, <)$? Is it derivable?

Solution. True in $(\mathbb{N}, <)$*?* No. The hypothesis $\forall x\,\exists y\,(x < y)$ is true in $(\mathbb{N}, <)$: each d has $d + 1$ above it. The conclusion $\exists y\,\forall x\,(x < y)$ is false: a candidate e fails at $x = e$, since $e < e$ is false. By (S5) of Definition 7.1.3, an implication with a true hypothesis and a false conclusion is false.

Derivable? No. By the soundness statement of Section 7.2, a derivable sentence is true in every structure; this one is false in $(\mathbb{N}, <)$, so no derivation exists.

Why these tools. A structure settles the question in two lines, while a derivation search could only fail, and a failed search proves nothing. A truth table is the wrong look-alike: it reads the sentence as $p \to q$ with two atoms, and a refuting row for $p \to q$ says nothing here, since

the sentence $\forall x\, P(x) \to \exists x\, P(x)$ has the same shape and is derivable (Proposition 5.5.1, then $\to$I). The soundness statement used here is proved in Section 7.3.

7.6 What First-Order Logic Can and Cannot Say

This section is optional, and nothing outside it depends on it. Every theorem in it is **stated without proof**, except where a short argument from a stated theorem is given and marked as such; its results carry no optional marker of their own, and the section as a whole is the unit a reader may skip.

Section 7.4 asked of a theory two questions that sound different: is it consistent, that is, does no finite list of its axioms derive $\bot$? And has it a model? A model answers the first question (Proposition 7.4.2), but nothing so far says that a consistent theory must have one. This section states the theorem that says so, Gödel's completeness theorem, and then follows its consequences. The first is compactness: a theory has a model as soon as each finite part of it has one. From compactness come the surprises. A first-order theory cannot fix the size of its infinite models; the true sentences about $\mathbb{N}$ have a model with an element larger than every number; and the property "finite" cannot be written as a sentence. The section ends with what can and cannot be decided by a mechanical procedure, and names the incompleteness theorem, which is a different matter from completeness.

Remark (Looking back). Section 4.8 proved that every valid propositional claim is derivable, and the optional Section 4.10 proved compactness for propositional logic. Both hold for predicate logic too; this section states them, and then shows what they cost: a first-order language cannot pin down the natural numbers.

7.6.1 Completeness

The soundness statement of Section 7.2 runs from derivations to truth: whatever a list of premises derives is true in every structure where the premises are true. The completeness theorem runs back.

Theorem 7.6.1. Gödel's Completeness Theorem

For every theory T and sentence σ: $T \vdash \sigma$ if and only if $T \models \sigma$. Equivalently, a theory is consistent if and only if it has a model.

This theorem is stated without proof. One direction is the soundness statement of Section 7.2; the other is new. Here $T \vdash \sigma$ and $T \models \sigma$ are those of Definition 7.4.1 and Definition 7.4.2: some finite list of axioms derives σ, and σ is true in every model of T. The two forms say the same thing. Take $\sigma = \bot$: $T \vdash \bot$ says that T is inconsistent (Definition 7.4.3), and $T \models \bot$ says that no model of T exists, since $\bot$ is true in no structure. In the other direction, $T \models \sigma$ means that the theory T with $\neg\sigma$ added, the counterexample set of the claim, has no model; and $T \vdash \sigma$ means that this theory is inconsistent, by $\neg$E in one direction and RAA in the other.

What the theorem means for this chapter is best said with its two halves side by side. Faced

with a claim $T \vdash \sigma$, Chapters 5 and 6 searched for a derivation, and this chapter searched for a counter-model, a model of T in which σ is false. Soundness says that both cannot succeed. Completeness says that one of them must: a stalled derivation search and a missing counter-model cannot both be real. The words "consistent" and "has a model", which Section 7.4 kept apart and never swapped, are now known to coincide, but only here, and with a theorem stated and not proved. Nothing outside this section uses the coincidence.

Section 7.4's group axioms show both outcomes. They derive that the identity is unique (Proposition 7.4.1), and that settles one claim by a derivation. They do not derive (Comm), and the triangle group of Example (The Six Symmetries of a Triangle), a model of the axioms in which (Comm) is false, settles that claim by a counter-model. Completeness says that every claim about groups is settled in one of these two ways, though it does not say which, or how long the search for either will take.

The theorem says nothing about how to find the derivation or the model, and the last part of this section states, without proof, that no method finds one of them in every case.

Second reading · *How the proof goes: a model made of names*

This is a sketch, not a proof. It follows the route of Section 4.10's proof of propositional compactness through Lindenbaum's lemma (Lemma 4.10.2).

Start from a consistent theory T; we want a model. The idea is to build the model out of the language itself, with the closed terms as its elements. Two difficulties arise, and each has a remedy.

Witnesses. A sentence $\exists x\, \varphi$ of the theory may hold with no closed term to show for it. So add new constants, and for each sentence $\exists x\, \varphi$ a **witness axiom** $\exists x\, \varphi \to \varphi[c/x]$, with c a constant new to everything so far. Adding it keeps the theory consistent, for the same reason the name in an $\exists$E box must be fresh: an old name already carries information, and a new one carries none, so assuming that it names a witness can lead to no contradiction that the bare $\exists x\, \varphi$ did not already give. This is repeated, since the new constants make new sentences, each with its own witness.

Deciding every sentence. A consistent theory may leave a sentence undecided, and a structure cannot. So extend the theory, one sentence at a time, to a **maximal** consistent set: list all sentences, and add each one if the result stays consistent. As in Lindenbaum's lemma, the result decides every sentence, containing it or its negation.

The structure. Its elements are the closed terms, with two terms identified when the set contains their equation; a constant names itself, a function symbol builds the longer term, and a predicate holds of the terms for which the set says it holds. An induction on sentences, the propositional one of Section 4.10 with a case for each quantifier, where the witness axioms do their work, shows that this structure satisfies exactly the sentences of the set. It satisfies T in particular.

The details, for countably many symbols, fill a chapter of a logic text; this book stops at the outline.

7.6.2 Compactness

Completeness ties models to derivations, and a derivation uses only finitely many premises. That one fact has a consequence about models alone.

Theorem 7.6.2. Compactness Theorem

If every finite subset of a theory T has a model, then T has a model.

A short argument from completeness. Suppose T had no model. Then T is inconsistent, by Theorem 7.6.1, so some finite list of its axioms derives $\bot$ (Definition 7.4.3). The finite subset made of those axioms is then itself an inconsistent theory, so by Proposition 7.4.2 it has no model, against the hypothesis.

Warning. "Every finite subset has a model" is not "has a finite model". The finite subsets are finite lists of **axioms**; the models that satisfy them may be infinite, and the model the theorem gives may be infinite even when every finite subset has a finite model. Compactness never makes a structure smaller.

Read the theorem from the side of what a theory forbids. A structure fails to be a model of T when some axiom is false in it. The theorem says that a property needing infinitely many axioms at once to fail is never enforced by first-order sentences: whatever a theory forbids, some finite part of it already forbids. That is the source of everything below.

The standard tool is a list of sentences that count. For $n \geq 2$, let σ_n be the sentence "there are at least n elements": n variables bound by $\exists$, followed by the conjunction of $\neg(x_i = x_j)$ for all $i < j$. It is true in a structure exactly when the domain has n or more elements. Any finite set of the σ_n is true in a large enough finite structure; all of them together are true only in an infinite one. Every use of compactness below follows one pattern: add infinitely many axioms that no single intended structure satisfies all at once, check that each finite part of the result has a model, and conclude that the whole has one. The model it gives is then a new structure, with properties the intended ones lack.

Quick check. Let σ_n say "there are at least n elements". Write σ_3.

Solution. $\exists x\, \exists y\, \exists z\, (\neg(x = y) \wedge \neg(x = z) \wedge \neg(y = z))$.

Quick check. Does compactness hold for propositional logic?

Solution. Yes: the optional Section 4.10 proved it; this section states the first-order version.

7.6.3 Large models and small ones

A set can be **listed** if its elements can be written as a sequence $d_0, d_1, d_2, \ldots$ in which every element appears; $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$ can, and so can every non-empty finite set (repeat its last element forever). "Listed" here has its everyday meaning.

Theorem 7.6.3. Löwenheim–Skolem Theorem

Let T be a theory in countably many symbols (the symbols can be listed $s_0, s_1, s_2, \ldots$). If T has an infinite model, then it has a model whose domain can be listed as $d_0, d_1, d_2, \ldots$, and also a model whose domain cannot be so listed.

This theorem is stated without proof. Two consequences show its force. First, take for T all the sentences true in $(\mathbb{R}; 0, 1, +, \cdot, <)$. The theorem gives a structure that satisfies every one of them and whose elements can be listed, although the elements of $\mathbb{R}$ cannot be (a fact stated here, not proved). Every first-order property of the real numbers holds in that structure, so the property that separates $\mathbb{R}$ from a listable set is not first-order. Second, take for T the sentences true in $(\mathbb{N}; 0, S, +, \cdot, <)$, with S the successor. The theorem gives a model of all of them that cannot be listed: far too large to be $\mathbb{N}$ under other names.

So no first-order theory with an infinite model fixes the size of its models. Section 7.5 showed that isomorphic structures satisfy the same sentences; here are structures that satisfy the same sentences and are not isomorphic, since one can be listed and the other cannot.

How can a listable structure satisfy every first-order sentence true in $\mathbb{R}$, when $\mathbb{R}$ is "complete": every non-empty set of reals bounded above has a supremum? The answer is the one the next two subsections give for $\mathbb{N}$. Completeness of $\mathbb{R}$ speaks of **all** sets of reals. A first-order sentence can speak only of sets defined by formulas (Section 7.5): for each formula there is a sentence "if the set this formula defines is non-empty and bounded above, it has a supremum", and each of these is true in $\mathbb{R}$. The listable model satisfies all of them, and fails the statement about all sets, which is not a first-order sentence. It must fail it: the listable model is an ordered field, as $\mathbb{R}$ is, and any two complete ordered fields are isomorphic (*Number Systems: From* $\mathbb{N}$ *to* $\mathbb{R}$, Chapter 6), so a complete listable model would make $\mathbb{R}$ listable, and $\mathbb{R}$ cannot be listed.

Quick check. A theory T in countably many symbols has the model $(\mathbb{Q}, <)$. Must it also have a model whose domain cannot be listed?

Solution. Yes: $(\mathbb{Q}, <)$ is an infinite model, so by Theorem 7.6.3 T also has a model that cannot be listed. For the strict order axioms, $(\mathbb{R}, <)$ is one.

7.6.4 Numbers that are not numerals

The symbols of arithmetic are 0, the one-place S for successor, $+$, $\cdot$ and $<$, read in $\mathbb{N}$ as usual. The **numeral** $\underline{n}$ is the closed term $S(S(\cdots S(0)\cdots))$ with n letters S; its value in $\mathbb{N}$ is n.

Example. A Model of Arithmetic with an Infinite Number

Let T be the set of all sentences in $0, S, +, \cdot, <$ that are true in $\mathbb{N}$. Add a new constant c,

and to T add the axioms

$$\underline{0} < c, \quad \underline{1} < c, \quad \underline{2} < c, \quad \ldots,$$

one for each $n \in \mathbb{N}$. Call the result T'.

Every finite subset of T' has a model. A finite subset contains finitely many sentences of T and finitely many axioms $\underline{n} < c$; let N be the largest n among them (or 0 if there is none). Read the symbols in $\mathbb{N}$ as usual and c as $N + 1$. The sentences of T do not mention c, and they are true in $\mathbb{N}$; each $\underline{n} < c$ in the subset says $n < N + 1$, which is true.

So T' has a model. By Theorem 7.6.2 there is a structure $\mathfrak{M}$ in which every sentence of T' is true. In $\mathfrak{M}$ every sentence true in $\mathbb{N}$ is true, and $c^{\mathfrak{M}}$ is above the value in $\mathfrak{M}$ of every numeral. It differs from each of them: if $c^{\mathfrak{M}}$ were the value of $\underline{n}$, the axiom $\underline{n} < c$ would put an element below itself, while $\forall x\, \neg(x < x)$ is true in $\mathbb{N}$, so in T, so in $\mathfrak{M}$.

An element of $\mathfrak{M}$ that is the value of no numeral is called **non-standard**; the values of numerals are the **standard** elements. $\mathfrak{M}$ satisfies every first-order sentence that $\mathbb{N}$ satisfies, and still has an element that no numeral names. Its standard elements lie at the bottom: the sentences $\forall x\, (x < \underline{n+1} \to x = \underline{0} \vee \cdots \vee x = \underline{n})$, one for each n, together with $\forall x\, \neg(x < \underline{0})$, are true in $\mathbb{N}$ and so in $\mathfrak{M}$, and they say that nothing non-standard lies below a standard element; since $\forall x\, \forall y\, (x < y \vee x = y \vee y < x)$ is also true in $\mathbb{N}$ and so in $\mathfrak{M}$, every standard element lies below every non-standard one.

The model $\mathfrak{M}$ is not $\mathbb{N}$ under another name. In $\mathbb{N}$ itself, read as usual, no value of c works for the whole of T': if c is read as the number n, the axiom $\underline{n} < c$ says $n < n$. Each finite part of T' is satisfied by $\mathbb{N}$, each with its own value for c; the whole set is satisfied only by a structure that compactness supplies, and that structure has elements $\mathbb{N}$ lacks. Since $\mathfrak{M}$ satisfies the same sentences in $0, S, +, \cdot, <$ as $\mathbb{N}$, no first-order sentence about arithmetic can tell the two apart; yet they are not isomorphic, since an isomorphism would send each numeral's value to the same numeral's value, and $c^{\mathfrak{M}}$ to an element of $\mathbb{N}$ above every number, and there is none.

Quick check. Is $c^{\mathfrak{M}} = S^{\mathfrak{M}}(d)$ for some element d of $\mathfrak{M}$?

Solution. Yes. The sentence $\forall x\, (\neg(x = 0) \to \exists y\, (S(y) = x))$ is true in $\mathbb{N}$, so it is true in $\mathfrak{M}$; and $c^{\mathfrak{M}} \neq 0^{\mathfrak{M}}$, since $\underline{0} < c$ is an axiom and $\forall x\, \neg(x < x)$ holds in $\mathfrak{M}$.

7.6.5 Induction: a schema and an axiom

How can $\mathfrak{M}$ satisfy induction, when its standard elements contain 0, are closed under successor, and are not everything? The answer turns on what first-order induction says.

The **induction schema** of first-order arithmetic is one sentence for each formula φ with free variable x:

$$(\varphi[0/x] \wedge \forall x\, (\varphi \to \varphi[S(x)/x])) \to \forall x\, \varphi,$$

with its other free variables, if any, bound by $\forall$ in front. Each instance is true in $\mathbb{N}$, by induction, so each is in T and true in the model $\mathfrak{M}$ of Example (A Model of Arithmetic with an Infinite Number). Yet the set of standard elements of $\mathfrak{M}$ contains $0^{\mathfrak{M}}$ and is closed under $S^{\mathfrak{M}}$, and is not all of the domain. There is no contradiction: the schema speaks only about sets **defined by formulas** (Section 7.5), with any other free variables fixed at elements of the domain, and

the standard elements are not such a set.

The argument is short. Suppose a formula φ, with free variable x and its other free variables $z_1, \ldots, z_k$ (if any) fixed at elements $e_1, \ldots, e_k$ of $\mathfrak{M}$, defined the standard elements in $\mathfrak{M}$; every assignment below sends each z_i to e_i. By Lemma 7.3.1 (Section 7.3; 0 is closed, and $S(x)$ is free for x in φ), $\varphi[0/x]$ holds in $\mathfrak{M}$ exactly when φ holds of $0^{\mathfrak{M}}$, and $\varphi[S(x)/x]$ holds under an assignment ℓ exactly when φ holds of $S^{\mathfrak{M}}(\ell(x))$. Then $\varphi[0/x]$ holds in $\mathfrak{M}$, since $0^{\mathfrak{M}}$ is standard, and $\forall x\,(\varphi \to \varphi[S(x)/x])$ holds, since the successor of a standard element is standard. The induction instance for φ is a sentence true in $\mathbb{N}$, so it is in T and true in $\mathfrak{M}$; read with each z_i at e_i, it gives $\forall x\,\varphi$ in $\mathfrak{M}$, so every element would be standard, against $c^{\mathfrak{M}}$. (Exercise C2 reaches the same conclusion another way, by least elements.)

The induction principle of Section 4.2 (Axiom 4.2.1) is stated for **every** property of natural numbers. Read in a structure, that is a statement about all subsets of the domain, not only the definable ones: it quantifies over sets, and is a **second-order** statement. Applied to the set of standard elements, it would rule out non-standard elements at once, and so a structure satisfying it, with 0 and successor read as usual, has no room for $c^{\mathfrak{M}}$. First-order logic can state the schema, not the axiom; compactness is the price of its completeness.

7.6.6 What can be decided

A question about sentences is **decidable** if a mechanical procedure answers it, correctly and in finitely many steps, for every sentence. Three facts, each stated without proof here, mark the boundary.

Remark. **Decidable: propositional logic.** Chapter 4's search tree decides every propositional claim in finitely many steps (Proposition 4.6.1): it either closes, and the claim is valid, or it leaves an open branch, which gives a refuting row.

Remark. **Undecidable: predicate logic (Church and Turing, 1936).** No mechanical procedure decides, for every first-order sentence, whether it is derivable. With completeness, the boundary can be said exactly: listing all derivations, shortest first, finds one when one exists; what no procedure can do is announce, in every case, that none exists. A stalled search, the situation of Chapter 6, is the normal state of affairs and not a sign that more cleverness was needed.

Remark. **Decidable: the monadic fragment.** For sentences with only one-place predicate symbols and no constants, function symbols or $=$, the bound of Section 7.2's remark makes the counter-model search finite: if such a sentence has a counter-model at all, it has one of bounded size, and there are finitely many structures of each size to try.

Remark. **Incompleteness, named.** Gödel showed in 1931 that no consistent theory whose axioms a mechanical procedure can list, and which derives the basic facts of arithmetic, derives every sentence true in $\mathbb{N}$. Completeness says every sentence true in **all** models of a theory is derivable from it; incompleteness says that the sentences true in **the** intended

model $\mathbb{N}$ are not all true in all models of any such theory. The two theorems do not conflict: the non-standard models are where they part.

Quick check. A sentence σ is true in $\mathbb{N}$ and false in some model of a theory T of arithmetic. Can $T \vdash \sigma$?

Solution. No: by the soundness statement of Section 7.2, $T \vdash \sigma$ would make σ true in every model of T.

Remark (Looking ahead). *Number Systems: From $\mathbb{N}$ to $\mathbb{R}$*, Chapter 1, takes the second-order induction principle as clause (P5) of a Peano system and proves that any two Peano systems are isomorphic; this section explains why no first-order list of axioms could do the same.

7.6.7 Exercises

7.6.7.1 A. Check your understanding

Exercise. A1

(a) State Theorem 7.6.1 in its two forms.
(b) State Theorem 7.6.2, and give the short argument from completeness in three sentences.
(c) What is a non-standard element of a model of arithmetic?
(d) Why does the induction schema not rule non-standard elements out?
(e) Name one decidable and one undecidable question about sentences.

Solution. *(a)* For every theory T and sentence σ, $T \vdash \sigma$ if and only if $T \models \sigma$. Equivalently, a theory is consistent if and only if it has a model.

(b) If every finite subset of T has a model, so has T. If T had no model, it would be inconsistent by completeness, so a finite list of its axioms would derive $\bot$. That finite subset would be inconsistent, and so, by Proposition 7.4.2, without a model, against the hypothesis.

(c) An element that is not the value of any numeral $\underline{n}$.

(d) Each instance speaks of a set defined by a formula, and the standard elements are not such a set (the argument in "Induction: a schema and an axiom").

(e) Decidable: validity of a propositional claim (or of a monadic sentence). Undecidable: derivability of a first-order sentence.

Exercise. A2: Drill

Each part takes a line or two. The parts run from writing a counting sentence, through

naming the theorem that answers a question, to judging what a theorem does and does not say.

(a) Write σ_2, "there are at least two elements".

(b) Write "there are at most two elements".

(c) Which theorem says "every consistent theory has a model"?

(d) Which says "if every finite part of T has a model, T has one"?

(e) True or false: completeness says that every sentence true in $(\mathbb{N}; 0, S, +, \cdot, <)$ is derivable with no premises.

(f) True or false: a theory with models of every finite size has an infinite model.

(g) Is "the domain is finite" expressible by one sentence?

(h) True or false: in a model of all sentences true in $\mathbb{N}$, every element is the value of a numeral.

(i) Over what does the second-order induction principle quantify that the schema does not?

(j) Is derivability of a propositional claim decidable? Of a monadic first-order sentence (one-place predicate symbols only; no constants, function symbols or $=$)? Of a first-order sentence in general?

(k) Does $(\mathbb{R}; 0, 1, +, \cdot, <)$ have a model of the same sentences whose domain can be listed?

(l) Why is a theory whose models are exactly the structures of size 3 possible, and one whose models are exactly the finite structures not?

Solution.

(a) $\exists x\, \exists y\, \neg(x = y)$.

(b) $\forall x\, \forall y\, \forall z\, (x = y \vee x = z \vee y = z)$.

(c) Completeness, in its second form.

(d) Compactness.

(e) False: with no premises, completeness makes derivable only the sentences true in **every** structure. $\forall x\, \neg(S(x) = 0)$ is true in $\mathbb{N}$ and false in a one-element structure (Section 7.2).

(f) True, by compactness, as in B1.

(g) No (C1).

(h) False: Example (A Model of Arithmetic with an Infinite Number) gives such a model with a non-standard element.

(i) All sets (properties) of elements, not only those defined by formulas.

(j) Yes; yes; no.

(k) Yes, by Theorem 7.6.3.

(l) "At least 3 and at most 3 elements" is one sentence, σ_3 with the "at most" sentence. No theory can do it for "finite": such a theory would have a model of every finite size (the structures built in the solution of C1), so by Exercise (B1: Arbitrarily large finite models) it would have an infinite model, which is not finite.

7.6.7.2 B. Practice

Exercise. B1: Arbitrarily large finite models

Prove: if a theory T has, for every n, a finite model with at least n elements, then T has an infinite model. Apply it to the theory of an equivalence relation R in which every class has exactly two elements, with that last condition written as one sentence.

Solution. Let T' be T with every σ_n, $n \geq 2$, added. A finite subset of T' contains only finitely many σ_n; let N be the largest such n (or 2 if there is none). By hypothesis T has a finite model with at least N elements, and it satisfies every σ_n with $n \leq N$, so it is a model of the finite subset. By Theorem 7.6.2 T' has a model. It satisfies every σ_n, so it has at least n elements for every n: it is infinite. And it is a model of T.

Application. The axioms are (Refl), (Sym), (Trans) for R and

$$\forall x\, \exists y\, (\neg(x = y) \wedge R(x, y) \wedge \forall z\, (R(x, z) \rightarrow (z = x \vee z = y))),$$

"each x is related to exactly one element besides itself". For each $k \geq 1$, a set of $2k$ elements split into k pairs, with R "in the same pair", is a model, so there are models with at least n elements for every n. So an infinite model exists; $\mathbb{Z}$ with $R(x, y)$ read as "x and y lie in the same pair $\{2m, 2m + 1\}$ for some $m \in \mathbb{Z}$" is one.

Marking scheme. (10 points) 3 for T' defined by adding axioms that force infinitely many elements (the σ_n, or distinct new constants c_i, the model then cut back to T's symbols). 3 for every finite subset having a model, with the size argument. 2 for compactness applied and the model shown infinite. 2 for the application's sentence and models. Fatal error (0 for the proof): applying compactness to T' without checking its finite subsets.

Exercise. B2: Which theorem settles it?

For each question, say which statement of this section or of Section 7.2 settles it (or that none does), and answer it.

(i) $\models \sigma$ is known. Is $\vdash \sigma$?

(ii) $T \nvdash \bot$. Has T a model?

(iii) σ is true in every finite structure. Is σ valid?

(iv) Is there a structure satisfying every sentence true in $\mathbb{N}$, with an element above the value of every numeral?

(v) Is there a procedure that decides validity for sentences built from one-place predicate symbols only (no constants, function symbols or $=$)?

Solution.

(i) Completeness: yes.

(ii) Completeness, second form: yes.

(iii) No theorem makes it so, and it need not be: the sentence of Exercise (C1: True in every finite model, used later) is true in every finite structure and false in $\mathbb{N}$ with f read as the successor, so it is not valid.

(iv) Compactness: yes (Example (A Model of Arithmetic with an Infinite Number)).

(v) The monadic remark: yes.

Marking scheme. (10 points) 2 for each part: 1 for the statement named (in (iii), for saying that no statement settles it), 1 for the answer (in (iii), "no", with a counter-model). Fatal error (0 for (iii)): answering "yes, by completeness".

7.6.7.3 C. Going deeper

Exercise. C1: "Finite" is not first-order

Prove that no sentence φ is true in exactly the finite structures for its symbols.

Solution (Hint 1). Use Exercise (B1: Arbitrarily large finite models).

Solution (Hint 2). Suppose φ were one, and consider the theory $T = \{\varphi\}$.

Solution. Suppose φ is true in every finite structure for its symbols and in no infinite one. For each $n \geq 1$ there is a structure of size n for the symbols of φ: take any n-element set, read each constant as one fixed element a of it, each function symbol as the function with the constant value a, and each predicate symbol as the empty relation. It is finite, so φ is true in it. So $T = \{\varphi\}$ has finite models of every size, and by Exercise (B1: Arbitrarily large finite models) it has an infinite model, in which φ is true: a contradiction.

Marking scheme. (10 points) 3 for the supposition stated precisely. 4 for models of every finite size exhibited. 3 for B1 (or compactness directly) applied, and the contradiction named. Fatal error (0 for the proof): a "proof" that only shows that one particular sentence fails.

Exercise. C2: Challenge: least elements in a non-standard model

Let $\mathfrak{M}$ be the model of Example (A Model of Arithmetic with an Infinite Number).

(a) For a formula ψ with free variable x, write a sentence saying "if some element satisfies ψ, some least element satisfies ψ", and explain why it is true in $\mathfrak{M}$.

(b) Take $\psi = \neg\varphi$, where φ is a formula supposed to define the standard elements in $\mathfrak{M}$. The sentence of (a) then names a least non-standard element e. Choose an element below e that is also non-standard, and conclude, without using the induction argument of this section, that no such φ exists.

Solution. (a) Choose a variable y that does not occur in ψ, free or bound; no quantifier of ψ is on y, so y is free for x in ψ (Definition 5.2.6). The sentence is

$$\exists x\, \psi \to \exists x\, (\psi \wedge \forall y\, (y < x \to \neg\psi[y/x])),$$

with any other free variables of ψ bound by $\forall$ in front. By Lemma 7.3.1 (Section 7.3; y is free for x), $\neg\psi[y/x]$ is satisfied under ℓ exactly when ψ fails of $\ell(y)$, so the sentence says what (a)

asks, in $\mathbb{N}$ and in $\mathfrak{M}$ alike. It is true in $\mathbb{N}$ by the least-number principle (every non-empty set of natural numbers has a least element), whatever set ψ defines there. So it belongs to T, and it is true in $\mathfrak{M}$.

(b) $\neg\varphi$ is satisfied by $c^{\mathfrak{M}}$, which is non-standard, so (a) gives a least e satisfying $\neg\varphi$: a least non-standard element. $e \neq 0^{\mathfrak{M}}$, since $0^{\mathfrak{M}}$ is standard; and $\forall x\,(\neg(x = 0) \to \exists y\,(S(y) = x))$ is true in $\mathbb{N}$, so in $\mathfrak{M}$. Choose d with $S^{\mathfrak{M}}(d) = e$. Then $d < e$ in $\mathfrak{M}$, since $\forall y\,(y < S(y))$ is true in $\mathbb{N}$. And d is non-standard: if d were the value of $\underline{n}$, then $e = S^{\mathfrak{M}}(d)$ would be the value of $\underline{n+1}$. So d satisfies $\neg\varphi$ and lies below e, against the choice of e.

Marking scheme. (10 points) 4 for (a): the sentence, with y chosen not to occur in ψ, so that it is free for x, and the reading of $\neg\psi[y/x]$ by Lemma 7.3.1, its truth in $\mathbb{N}$ by the least-number principle, and the transfer to $\mathfrak{M}$ through T. 4 for (b): the predecessor d obtained from a sentence true in $\mathbb{N}$, with $d < e$ and d non-standard both argued. 2 for the contradiction named. Fatal error (0 for (a)): the least-number principle used for an arbitrary subset of $\mathfrak{M}$ rather than through a sentence of T.

Chapter 8

From Derivations to Mathematical Prose

Here is a proof that every schoolbook prints: for an integer n, if n is even, then n^2 is even.

> Suppose n is even, say $n = 2k$. Then $n^2 = 4k^2 = 2(2k^2)$, so n^2 is even.

And here is the same proof as a derivation, for one fixed integer n, with "even" written out: n is even when $n = 2k$ for some integer k.

1	$\exists k\,(n = 2 \cdot k)$	asm, for $\to$I
2	k_0 | $n = 2 \cdot k_0$	asm, for $\exists$E
3	$n \cdot n = 2 \cdot (2 \cdot k_0 \cdot k_0)$	arith 2
4	$\exists j\,(n \cdot n = 2 \cdot j)$	$\exists$I 3
5	$\exists j\,(n \cdot n = 2 \cdot j)$	$\exists$E 1, 2–4
6	$\exists k\,(n = 2 \cdot k) \to \exists j\,(n \cdot n = 2 \cdot j)$	$\to$I 1–5

At line 5 the side condition holds: k_0 is not in line 5, not in line 1, and in no other open assumption. The words and the lines match one for one:

Words	Lines	What they do
"Suppose n is even"	1	open the $\to$I box
"say $n = 2k$"	2	open the $\exists$E box, with a new name
"Then $n^2 = 4k^2 = 2(2k^2)$"	3	algebra inside both boxes
"so n^2 is even"	4	$\exists$I, with the witness $2k^2$
(the paragraph ends)	5, 6	the two boxes close

Words such as "let", "choose", "suppose" and "in either case" are the boxes, spoken. This chapter makes that correspondence exact in both directions. Learning to hear the boxes is what lets a reader working alone check a proof, by restoring the derivation underneath wherever a step is in doubt; learning to speak them is what makes a proof checkable by someone else.

Words can also deceive. Here is a proof whose every sentence is true and which is still wrong.

Claim. For an integer n, if n^2 is even, then n is even.

(1) Suppose n is even. (2) Then $n = 2k$ for some integer k. (3) So $n^2 = 4k^2 = 2(2k^2)$. (4) Hence n^2 is even. (5) Therefore, if n^2 is even, then n is even.

Every sentence is true, and so is the claim. But the box opens with "n is even", so what the lines prove is the converse, "if n is even, then n^2 is even". Section 8.6 checks this proof line by line; by the end of the chapter you will see the error at a glance.

The question of this chapter: when a proof is written in sentences, where did its boxes go, and how does a reader find them again?

The chapter builds a dictionary from rules to phrases, writes proofs at four levels from full derivation to plain prose, fixes which steps plain prose may leave unsaid, and turns all of this into habits for reading, writing and checking proofs, with a catalogue of the errors that recur. It ends with the one rule mathematicians have seriously argued about: reductio ad absurdum.

What you need. Chapters 1–6, core. From Chapter 7, only the statement in Section 7.2 under "From counter-models to non-derivability", that a derivation with the quantifier rules never leads from true premises to a false conclusion; Sections 8.6 and 8.7, and Exercise C3 of Section 8.5, use it as a fact.

- **Boxes.** An assumption opens a box, the line written after it closes it, and a line may cite only lines and boxes available to it (Definition 1.4.2).
- **The rules.** The introduction and elimination rules of every connective and quantifier, and the derived rules of Chapter 2; Section 8.1 collects them on one card.
- **Choosing.** "Choose" needs a fresh name, and a chosen constant may depend only on the names of the boxes around it (Sections 5.6 and 6.4).
- **The classical rule.** RAA is the one rule that belongs to no connective; ¬¬E and LEM are its other two forms (Section 2.5).

↻ **Second reading** · *The chief earlier results this chapter cites*

From Chapters 1 and 2: the rules for every connective, derivations with their boxes and scope rule (Definition 1.4.2), derived rules (Definition 2.3.1), the classical rules of Section 2.5 and their interderivability (Theorem 2.5.7), the cut rule (Lemma 2.3.3), and the two running examples Theorem 2.1.3 and Theorem 2.2.1. From Chapter 3: the search for a derivation (Section 3.1), the extraction of an argument (Section 3.3) and Proposition 3.2.2. From Chapter 4: induction and strong induction (Axiom 4.2.1, Theorem 4.2.2), truth tables and refuting rows (Example (Working backwards from F), Definition 4.5.3, Corollary 4.7.3), soundness (Theorem 4.7.2), completeness (Theorem 4.8.4, Corollary 4.8.5), the rule TAUT (Definition 4.8.1) and the proof of Lemma 4.8.3 (for a Challenge exercise of Section 8.4), and, for Section 8.7 only, the optional Theorem 4.7.4. From Chapters 5 and 6: the quantifier rules and their side conditions (Definition 5.4.1, Definition 5.6.1), the bounded rules (Proposition 6.1.1), definitions of analysis translated into formulas (Section 6.1), the laws for negating quantifiers (Proposition 6.2.1), the propositional skeleton (Definition 6.3.1, Proposition 6.3.1), and the derivation that uniform continuity implies continuity together with the place where its converse breaks down (Example (Uniformly continuous implies continuous), Example (The search from C to UC), and the counter-model Example (The square function on the real line)). School arithmetic and the order of $\mathbb{R}$ are used in examples as facts. From Chapter 7: the statement of Section 7.2 named above (its proof is the optional Section 7.3). For the optional Section 8.8: structural induction and recursion (Theorem 4.3.1, Theorem 4.3.4) and the proof of Theorem 4.7.2 with its context

lemma (Lemma 4.7.1).

Time. About 35 hours on the core track (all of A, one B and one C per section, the exit checks); about 65 hours for everything.

Minimum before *Sets, Relations, and Cardinality*. Chapters 1–6 and 8, core; Sections 7.1–7.2 before its Chapter 3.

8.0.1 Roadmap

- **The connective rules in prose.** The rule card for the whole chapter, and which steps may go unsaid; the dictionary rows for the connectives and the classical rules, each with a short proof annotated phrase by phrase; the spoken purpose of a box; the constructive dilemma and equivalent conditions proved by a cycle; why ¬I and RAA are told apart; the three faces of "proof by contradiction"; the costumes.
- **The quantifier rules in prose.** The rows for "for every", "there exists", equality and unique existence; definitions, "the" and double inclusion; what a chosen constant depends on, in a picture and in numbers; "it suffices to show", induction and "without loss of generality"; the quantifiers that English leaves unwritten or puts in the wrong place.
- **Four levels of one proof.** One short proof written four times; the levels named; the convention on silent steps; the irrationality of $\sqrt{2}$ and "uniformly continuous implies continuous" at the levels that show most; which silent steps are truth-functional; "it suffices to show" written as a column; how to check a doubtful sentence by expanding it.
- **Reading a proof: self-explanation.** The book's training passage for the habit of naming, at each line, what makes it true; four questions to ask of every line; naming the unstated fact; and five strategies that good readers apply to a proof as a whole.
- **Writing a proof.** A seven-step procedure, run at once on one example; scratch paper against the page; the conventions of written mathematics; two more examples; generalizing a proof.
- **Checking proofs, and what is not allowed.** Two questions for every line; the catalogue (N1)–(N14) in seven families; four student proofs checked line by line (and, in a fold, the errors' traditional names).
- **The dispute over excluded middle.** A proof of existence that names nothing; what a proof provides under the BHK reading; reductio as the one extra rule; why excluded middle cannot be refuted; which laws survive without the rule; a short history; what analysis would lose; proofs as programs; the book's position.
- **Stages of knowledge: Kripke models** (optional). A semantics in which knowledge grows by stages; the rules without RAA are sound for it, and two or three stages show which laws need RAA.

8.0.2 Named moves

- **Hear the rule in the phrase.** "Let $x \in A$" opens a ∀I box and "suppose" an assumption; "choose" is ∃E; "take" and "let $\delta := \ldots$" are ∃I; "suppose, for contradiction" is RAA or ¬I; both boxes end in ⊥, and what the box assumes and the line written after it closes tell which.
- **Expand to check.** When a sentence is doubtful, rewrite that stretch, and only that stretch, as a skeleton, and look for the line that breaks a rule.
- **Skeleton first.** Write the first and the last sentence of every box before filling any.
- **Self-explain each line** by earlier lines, definitions and named results, and the rule that

combines them; never by paraphrase.
- **Two columns when checking:** *is the line true?* and *does it follow, and by which rule?* The first column finds where to look; the second decides.
- **Say what depends on what.** Every chosen object carries its dependencies: the names in the line it comes from, among the boxes open when it was chosen.
- **Write the reductions as a column.** Each "it suffices to show" is one row with its reason; read upwards, the column is the proof. (Section 8.3)
- **Generalize by a fresh name.** Replace the particular object by a new letter, and keep only the lines that still hold for it. (Section 8.5)

8.0.3 Summary

A written proof is a derivation in sentences: every rule has its phrases, and "let", "suppose", "choose" and "in either case" are the boxes, spoken (Sections 8.1–8.2). The chapter named four levels of one proof (Definition 8.3.1) and fixed by one principle which steps plain prose may leave unsaid (Definition 8.3.2); a doubtful sentence is checked by expanding it to a skeleton. It gave a habit for reading (self-explanation), a procedure for writing, and, for checking, two questions and a catalogue of fourteen errors (Example (What is not allowed)). It ended with the one rule mathematicians have disputed. Without RAA, excluded middle has no derivation (the optional Theorem 4.7.4) and yet cannot be refuted (Proposition 8.7.3). An irrational power of an irrational number can be shown rational with or without a named witness (Theorem 8.7.1, Proposition 8.7.2). The optional Section 8.8 proves, with stages of knowledge, which classical laws need the rule. What is open is what the logic is for. *Sets, Relations, and Cardinality* writes its proofs at levels 2 and 3 with these tools, and its first question is what a set is.

8.0.4 Terms

- **BHK reading**: a proof as a construction of what it claims (Section 8.7).
- **costume**: a direct or contrapositive proof dressed as contradiction (Example (A direct proof in costume)).
- **double inclusion**: $A = B$ shown as $A \subseteq B$ and $B \subseteq A$ (Example (Two inclusions)).
- **forcing** (optional): what a stage of knowledge makes true (Definition 8.8.2).
- **heredity** (optional): what a stage forces, every later stage forces (Lemma 8.8.1).
- **intuitionistic logic**: the rules of this book without RAA (Section 8.7).
- **Kripke model** (optional): stages of knowledge, with atoms that stay known (Definition 8.8.1).
- **levels of proof**: derivation, skeleton, annotated prose, plain prose (Definition 8.3.1).
- **not-allowed catalogue**, (N1)–(N14): recurring errors, each with the rule broken (Example (What is not allowed)).
- **self-explanation**: naming each line's warrant, never paraphrasing it (Example (Explaining a proof, well and badly)).
- **silent step**: a step plain prose may leave unwritten (Definition 8.3.2).
- **skeleton**: level 1, the boxes and the logical lines (Definition 8.3.1).
- **TFAE by a cycle**: several conditions shown equivalent by a loop of implications (Example (Three conditions, one cycle)).

8.1 The Connective Rules in Prose

Chapters 1–7 built a calculus in which every step of a proof is one rule applied to lines that are available. Mathematicians do not write Fitch boxes; yet every proof they write is a derivation

underneath, written in sentences. This section and the next compile the dictionary between the two: for each rule, the phrases that carry it, a sentence in which it does its work, and whether a writer may leave it unsaid. This section covers the connectives and the classical rules; Section 8.2 the quantifiers, equality and induction.

8.1.1 The rule card

Here is every primitive rule of the book, with the classical and derived rules most used in prose, on one card: its shape, the words that carry it in a written proof, and whether plain prose may leave it out. The last column is explained in the next subsection; Section 8.2 gives the quantifier rows their examples.

Rule	Shape	You write	Silent in plain prose?
$\wedge$I	$\varphi, \psi \Longrightarrow \varphi \wedge \psi$	"…, and …"; "both hold"	when both parts are displayed
$\wedge$E	$\varphi \wedge \psi \Longrightarrow \varphi$ (or ψ)	"in particular"	yes
$\rightarrow$I	box φ … $\psi \Longrightarrow \varphi \rightarrow \psi$	"Suppose P. … So Q."	the opening never; the closing yes
$\rightarrow$E	$\varphi \rightarrow \psi, \varphi \Longrightarrow \psi$	"since …"; "by (the lemma), …"	yes when the implication is a line of the proof itself; never when it is a hypothesis or a cited result
R	φ (available) $\Longrightarrow \varphi$, copied	"recall that"	yes
$\vee$I	$\varphi \Longrightarrow \varphi \vee \psi$ (or $\psi \vee \varphi$)	"so certainly … or …"	yes
$\vee$E	$\varphi \vee \psi$, box φ … χ, box ψ … $\chi \Longrightarrow \chi$	"Case 1 … Case 2 … In either case"	the split and why it is exhaustive never; "in either case" yes
$\neg$I	box φ … $\bot \Longrightarrow \neg\varphi$	"Suppose P. … a contradiction. Hence not P."	never
$\neg$E	$\varphi, \neg\varphi \Longrightarrow \bot$	"this contradicts …"	never: both lines of the clash are named
$\bot$E	$\bot \Longrightarrow \varphi$	"this case cannot occur"	yes, once the clash or the empty case is stated
$\leftrightarrow$I	box φ … ψ, box ψ … $\varphi \Longrightarrow \varphi \leftrightarrow \psi$	"($\Rightarrow$) … ($\Leftarrow$) …"; "conversely"	never either direction
$\leftrightarrow$E	$\varphi \leftrightarrow \psi, \varphi \Longrightarrow \psi$	"that is"; "by definition"	yes, for a definition
RAA	box $\neg\varphi$ … $\bot \Longrightarrow \varphi$	"suppose, for a contradiction, that not P"	never
$\neg\neg$E, LEM	$\neg\neg\varphi \Longrightarrow \varphi$; $\Longrightarrow \varphi \vee \neg\varphi$	"P must hold, since …"; "either P or not"	never: written out

Rule	Shape	You write	Silent in plain prose?
MT, CP	$\varphi \to \psi,\ \neg\psi \Longrightarrow \neg\varphi$; $\varphi \to \psi \Longrightarrow \neg\psi \to \neg\varphi$	"so P cannot hold"; "we prove the contrapositive"	the announcement never
DS	$\varphi \vee \psi,\ \neg\varphi \Longrightarrow \psi$	"the first case is impossible, so …"	the name yes; the fact it uses never
$\forall$I	box (fresh a) … $\varphi(a) \Longrightarrow \forall x\,\varphi(x)$	"let $x \in A$"; "fix $\varepsilon > 0$"	the opening never; "since x was arbitrary" yes
$\forall$E	$\forall x\,\varphi(x) \Longrightarrow \varphi(t)$	"in particular, for $x = \ldots$"	at the obvious term yes; otherwise the term is named
$\exists$I	$\varphi(t) \Longrightarrow \exists x\,\varphi(x)$	"take $N \coloneqq \ldots$"; "… will do"	never: the witness and what it depends on
$\exists$E	$\exists x\,\varphi(x)$, box (fresh a) $\varphi(a) \ldots\ \chi \Longrightarrow \chi$	"choose N such that …"	the choice with its fresh name never; the closing yes
=I, =E	$\Longrightarrow t = t$; $s = t,\ \varphi(s) \Longrightarrow \varphi(t)$	"… $=$ … $=$ …"; "substituting"	yes, inside chains

The exact conditions are those of Definition 8.3.2 in Section 8.3; the card abbreviates them.

Quick check. Which rule does each phrase carry: (i) "say $n = 2k$", after "n is even"; (ii) "in either case"; (iii) "this contradicts $x^2 \geq 0$"?

Solution.

(i) $\exists$E, opening the box of k ("n is even" unwound to $\exists k\,(n = 2k)$). (ii) $\vee$E, closing. (iii) $\neg$E, with both lines of the clash named.

The purpose of a box, spoken. Since Section 1.4 every derivation in this book has printed, beside an assumption, the rule that will close its box: "asm, for $\to$I", "asm, for $\neg$I", and so on. Prose says the same thing with the sentence that opens the box. "Suppose P" is asm, for $\to$I. "Suppose P; … this is impossible", or "if P were true", is asm, for $\neg$I. "Suppose, for a contradiction, that not P" is asm, for RAA. "Case 1: P" is asm, for $\vee$E. "Choose a such that $P(a)$" is asm, for $\exists$E. "($\Rightarrow$) Suppose P" is asm, for $\leftrightarrow$I. "Let $x \in A$" is asm, for $\forall$I, its bounded form. A box whose purpose the reader cannot hear is a box the writer did not open in words.

8.1.2 Which steps may go unsaid

Plain prose leaves steps out. If it left out whatever a writer found obvious, a reader could never tell an omission from a gap, so the book fixes the omissions by one principle: *a step may go unsaid when a reader who knows the rules can restore it uniquely from what is written.* A step that **creates** something (a name, a witness, a case split), **assumes** something, or **cites** something cannot be restored that way, so it is always written.

Silent steps are the ones whose result is forced by the lines around them; every choice, assumption and citation is spoken.

The last column of the card is this principle applied rule by rule.

Quick check. In "Since $0 < x < 1$, in particular $x > 0$", may the words "in particular $x > 0$" be left out if the next sentence uses $x > 0$? May "choose k with $n = 2k + 1$" be left out?

Solution. The first, yes: $\wedge$E has two possible results, and the next sentence shows which one was used, so the step is forced. The second, no: $\exists$E creates a name, and a name cannot be restored.

8.1.3 The same proof, twice

Section 2.3 proved Theorem 2.1.3 ("if n^2 is even, then n is even") a second way, in Example (Even square, even number, by contraposition), with the derived rules CP and DS. Here are four sentences for it, with the box of the second sentence spelled out:

> Suppose n^2 is even. If n^2 were odd, it would be both even and odd, which (E2) forbids; so n^2 is not odd. By (E3), read contrapositively, n is not odd. Since n is even or odd, by (E1), n is even.

Before reading on: where in the paragraph does a box open, and where does one close?

The paragraph accounts for every line of the derivation, though some lines have no words of their own.

Lines	Words that carry them	Rule
4	"Suppose n^2 is even."	$\to$I, the box opens
5	"If n^2 were odd"	$\neg$I, the box opens
6	"it would be both even and odd"	$\wedge$I
7	"which (E2) forbids"	$\neg$E, citing premise 2
8	"so n^2 is not odd"	$\neg$I, the box closes
9	"By (E3), read contrapositively"	CP, citing premise 3
10	"n is not odd"	$\to$E
11–12	"Since n is even or odd, by (E1), n is even."	DS, citing premise 1
13	nothing	$\to$I, the box closes

The derivation the table refers to is the one of Section 2.3, with the facts (E1)–(E3) of Section 2.1 (every integer is even or odd; no integer is both; the square of an odd integer is odd) as its premises, written here with its atoms in words:

1	n even $\vee$ n odd	
2	$\neg(n^2$ even $\wedge$ n^2 odd$)$	
3	n odd $\to$ n^2 odd	
4	$\quad n^2$ even	asm, for $\to$I
5	$\quad\quad n^2$ odd	asm, for $\neg$I
6	$\quad\quad n^2$ even $\wedge$ n^2 odd	$\wedge$I 4, 5
7	$\quad\quad \bot$	$\neg$E 6, 2
8	$\quad \neg n^2$ odd	$\neg$I 5–7
9	$\quad \neg n^2$ odd $\to$ $\neg n$ odd	CP 3
10	$\quad \neg n$ odd	$\to$E 9, 8
11	$\quad n$ odd $\vee$ n even	or commutes 1
12	$\quad n$ even	DS 11, 10
13	n^2 even $\to$ n even	$\to$I 4–12

Three features of the table recur through the whole dictionary. The **openings** of boxes always have words, which name what is assumed: "suppose", "if ... were". Some lines have **no words**: line 13 closes the box of line 4, and the paragraph simply stops; line 11 only reorders the disjuncts of (E1) for the schema of Proposition 2.3.4, and prose does not see the order of a disjunction (the principle above lets the exchange be silent: the order a schema wants is forced). And a **premise** is never a sentence of its own: it enters through its name, "(E2)" or "by (E1)", at the step that uses it.

Quick check. Which word of the paragraph tells the reader that line 5 is an assumption and not a claim, and what would the sentence assert if that word were replaced by "is"?

Solution. The word "were" (with "if"): "if n^2 were odd" supposes something in order to refute it, which is the opening of a $\neg$I box. With "is", the sentence "n^2 is odd, so it is both even and odd" would assert "n^2 is odd" as a claim rather than an assumption: it would sit, unjustified, directly in the box of line 4, where it is in fact false (n^2 is even there), and the proof would be wrong.

8.1.4 How to read the dictionary

Each group of rows below repeats the card's rows for one connective, with the phrases a reader must recognize, and then works one short proof annotated phrase by phrase. A tag such as [$\wedge$E] after a phrase says which rule it carries; any wording that does the same job is correct.

8.1.5 Conjunction

Rule	Phrases	Silent in plain prose?
$\wedge$I	"…, and …"; "we show (a) … and (b) …"; "both … hold"	often, when both conjuncts are displayed
$\wedge$E	"in particular"; "so $a > 0$"; using one clause of a hypothesis without comment	yes

In a proof, $\wedge$I is a claim with two parts, proved one after the other and then put together; often the parts are labeled (a) and (b), and the last sentence says nothing more. $\wedge$E is the step from a chain to one of its links: "since $x \in (0,1)$, in particular $x > 0$" takes the left conjunct of $0 < x \wedge x < 1$.

Example. Two parts, and one part

Let x be a real number with $0 < x < 1$. Prove that $x^2 < x$ and $x^2 > 0$. Tag the steps that carry the logic.

Solution. The hypothesis $0 < x < 1$ is the conjunction "$0 < x$ and $x < 1$"; so $x > 0$ [$\wedge$E] and $x < 1$ [$\wedge$E]. *(a)* Multiplying $x < 1$ by the positive number x gives $x^2 < x$ [algebra]. *(b)* As $x > 0$, $x^2 = x \cdot x > 0$ [algebra]. So $x^2 < x$ and $x^2 > 0$ [$\wedge$I].

Both $\wedge$E steps could be silent ("multiplying $x < 1$ by $x > 0$ gives …"), because the next sentence shows which part of the hypothesis is used. The final $\wedge$I may be silent too, since both parts are displayed: a proof of a two-part claim usually proves (a), proves (b), and stops.

8.1.6 Implication

Rule	Phrases	Silent in plain prose?
$\to$I	"assume …; then …; hence if … then …"; "suppose …"; "if …, then …" as the frame of a whole proof	never the assumption; the closing "hence" may be silent
$\to$E	"since …, …"; "by (hypothesis, lemma), …"; "hence"	the citation, never
R	"recall that"; "as noted above"	yes
MT, CP	"…; so P cannot hold"; "we prove the contrapositive"; "equivalently, if not Q then not P"	the announcement "we prove the contrapositive", never

In analysis, $\to$I is a proof that begins "Suppose $|x - 1| < 1$" and ends "so $x > 0$"; the example below writes it out. $\to$E is every citation of a theorem: "since 42 is divisible by 2 and by 3, it

is divisible by 6" applies the theorem used in Section 1.3 to two facts in hand, and the word "since" introduces those facts. For CP, Section 2.3's Example (Even square, even number, by contraposition) is the model, as the opening of this section showed. (Theorem 2.1.3 itself is proved by cases, with no contraposition.)

Example. A hypothesis assumed, a lemma cited, a contrapositive announced

Let x be a real number. (a) Prove that if $|x-1|<1$, then $x>0$. (b) Deduce that if $x \leq 0$, then $|x-1| \geq 1$. Tag each step with its rule.

Solution. We use the school fact that, for real u and r, $|u|<r$ holds exactly when $-r<u<r$.

(a) Suppose $|x-1|<1$ [assumption for $\to$I]. By the school fact, $-1<x-1<1$ [$\leftrightarrow$E], in particular $-1<x-1$ [$\wedge$E], so $x>0$ [arith].

(b) Suppose $x \leq 0$ [assumption for $\to$I]. Recall from (a) that $|x-1|<1$ implies $x>0$ [R]. Since $x \leq 0$, that is, $x>0$ fails, $|x-1|<1$ cannot hold [MT], so $|x-1| \geq 1$.

Part (a) ends at "$x>0$": the closing line of $\to$I, "hence, if $|x-1|<1$, then $x>0$", is silent. The assumption could not be: without "suppose", the sentence "$|x-1|<1$" would be a claim about every real x, and a false one.

In (b), "recall" is R, and it could be dropped. "Cannot hold" is MT, with P the statement $|x-1|<1$ and Q the statement $x>0$; the last step reads $\neg(|x-1|<1)$ as $|x-1| \geq 1$, by the order fact of Section 2.1. An equivalent (b) announces the derived rule instead: "This is the contrapositive of (a)", which is CP applied to (a), followed by $\to$E. That announcement may never be silent: a reader who sees "suppose $x \leq 0$" and then a citation of (a) must be told that (a) is being used backwards.

8.1.7 Disjunction and cases

Rule	Phrases	Silent in plain prose?
$\vee$I	"…, so certainly P or Q"; "in particular $x \neq 0$ or …"	yes
$\vee$E	"we consider two cases"; "Case 1 … Case 2 …"; "in either case"; "otherwise"	never the split; the "in either case" may be silent
DS	"since $x \neq 0$, the first case is impossible, so …"	the rule's name, yes; the fact $x \neq 0$ it uses, never
constructive dilemma (Example (Constructive dilemma); the derived rule CD of Section 2.3, Exercise (B4: Dilemmas as derived rules, used later))	"if P then R, and if Q then S; since P or Q, we get R or S"; "in the first case …, in the second …, so R or S"	never the split or the two implications; the closing "so R or S" yes
"or" and "and" commute (Example (First derivations with "or"), Proposition 1.2.1)	"P or Q" read as "Q or P", with no words of its own	yes

In analysis, $\vee$I hides inside every weakening of a strict inequality: "$x > 0$, so $x \geq 0$". $\vee$E is the proof of the triangle inequality in Section 1.6 (Example (The triangle inequality by cases)), which splits on the sign of $x + y$ and reaches $|x + y| \leq |x| + |y|$ in both cases. DS is the step "$x^2 = x$, so $x = 0$ or $x = 1$; since $x \neq 0$, $x = 1$".

Example. Two cases, a weakening, a refuted disjunct

Let x be a real number. (a) Prove that $|x| + x \geq 0$. (b) Prove that if $x^2 = x$ and $x \neq 0$, then $x = 1$. Tag each step with its rule.

Solution. (a) Every real number satisfies $x \geq 0$ or $x < 0$, and we consider the two cases [$\vee$E].

Case $x \geq 0$. Then $|x| = x$, so $|x| + x = 2x \geq 0$.

Case $x < 0$. Then $|x| = -x$, so $|x| + x = 0$, and certainly $|x| + x \geq 0$ [$\vee$I].

So $|x| + x \geq 0$ in either case [$\vee$E, closing].

(b) Suppose $x^2 = x$ and $x \neq 0$ [assumption for $\to$I]. Then $x(x - 1) = x^2 - x = 0$, and a product of two real numbers is 0 only if one factor is, so $x = 0$ or $x - 1 = 0$ [$\to$E]. Since $x \neq 0$, the first case is impossible, so $x - 1 = 0$ [DS], that is, $x = 1$.

In (a), the split was announced with the disjunction that makes it exhaust; $\vee$E needs that disjunction on a line, and a split without it is the missing-case error that Section 8.6 catalogues. The second case weakened $|x| + x = 0$ to $|x| + x \geq 0$, which is $\vee$I. The closing "in either case" could be silent. In (b), the words "the first case is impossible" carry the fact $x \neq 0$ that DS uses: the rule's name is not needed, but that fact is.

Warning. "Clearly" often hides a case split. "Clearly $\max\{a, b\} \geq (a + b)/2$" is true for all real a, b, but its proof is $\vee$E on "$a \leq b$ or $b \leq a$": if $a \leq b$, then $\max\{a, b\} = b = (b + b)/2 \geq (a + b)/2$, and the other case is the same with the letters exchanged. The case split is never silent; write it, or write "by cases on whether $a \leq b$" and the one computation.

8.1.8 Negation and contradiction

Rule	Phrases	Silent in plain prose?
$\neg$I	"suppose P; … a contradiction; hence $\neg P$"; "P is impossible, since …"	never
$\neg$E	"this contradicts …"; "which is impossible"; "but …"	never: the clashing pair is named
$\bot$E	"this case cannot occur"; "vacuously"; "there is nothing to prove"	yes, once the contradiction or the empty case is stated
DM, quantifier negation	"not all … so some … fails"; "it is not the case that …, so …"	the negated form must be written

In analysis, ¬I is the proof of Theorem 2.2.1: "suppose some rational r has $r^2 = 2$; ...; this is a contradiction; hence no rational number has square 2". ¬E is a sentence such as "then $n \leq 1/\varepsilon < n$, a contradiction": the chain gives $n < n$, which clashes with the fact that no number is less than itself. One strict link is essential, since $n \leq 1/\varepsilon \leq n$ gives only $n \leq n$, which clashes with nothing. ⊥E is "if $A = \emptyset$, there is nothing to prove", said of a claim about every element of A: with no element to check, the claim is vacuously true (Section 5.4).

The last row needs more room. To negate "s is an upper bound of A", that is, $\forall a \in A\ (a \leq s)$, a proof writes the negated form: **some** $a \in A$ has $a > s$. Write the negated form out; do not leave the reader to push a "not" through a quantifier in their head.

Second reading · *For the record: where the classical step sits*

The passage uses Proposition 6.2.3 (a), derived with RAA, and the order fact that $\neg(a \leq s)$ says $a > s$. A proof that $\mathbb{N}$ is not bounded above is the instructive case. It opens "suppose, for contradiction, that $\mathbb{N}$ is bounded above" and closes with "therefore $\mathbb{N}$ is not bounded above"; that is ¬I, despite the phrase. Its one classical step comes after the box: from "not every $n \in \mathbb{N}$ has $n \leq x$" it passes to "some $n \in \mathbb{N}$ has $n > x$", by Proposition 6.2.3 (a), which Section 6.2 derives with RAA through Proposition 6.2.1 (a) and Proposition 6.2.2, and then by trichotomy for $\neg(n \leq x)$, giving $n > x$. That step produces an n without naming it, and the proof needs no name.

Example. A negative claim, its positive form, and an empty case

(a) Prove that no real number is an upper bound of $\mathbb{R}$. (b) Prove the positive form: for every real s there is a real x with $x > s$. (c) Let A be an empty set of real numbers. Prove that every real number is an upper bound of A. Tag each step with its rule.

Solution. (a) Let s be a real number [∀I], and suppose s is an upper bound of $\mathbb{R}$ [assumption for ¬I]. Since $s+1$ is a real number, $s+1 \leq s$ [∀E at $s+1$], that is, $1 \leq 0$. This contradicts $0 < 1$ [¬E]. Hence s is not an upper bound of $\mathbb{R}$ [¬I].

(b) Let s be a real number [∀I]. Take $x \coloneqq s+1$ [∃I]; then $x - s = 1 > 0$, so $x > s$.

(c) Let s be a real number and $a \in A$ [∀I, →I]. Since A has no elements, $a \in A$ is impossible [¬E, against "$a \notin A$"], so there is nothing to prove: $a \leq s$ [⊥E]. Hence s is an upper bound of A.

Claim (a) is $\forall s\, \neg\forall x\ (x \leq s)$, and its proof writes the ¬I box in full: the assumption, the named clash, the conclusion. The positive form (b), $\forall s\, \exists x\ (x > s)$, says more: it hands over the number that beats s. Given (b) as a premise, (a) follows with no classical step: if s were an upper bound of $\mathbb{R}$, the x that (b) provides for s would satisfy $x > s$ and $x \leq s$. The general passage from the shape of (a) to the shape of (b), with no witness in hand, goes through Proposition 6.2.1 (a), derived with RAA, and trichotomy. So when both forms are within reach, prove the positive one. In (c), once the clash and the empty set are named, the ⊥E line itself may be silent. Part (c) is the prose form of the derivation in Section 5.4 that concludes $\forall u\, \forall x\ (x \in A \to x \leq u)$ from $\forall x\, \neg(x \in A)$, with its vacuous step unpacked into the ¬E and ⊥E lines.

8.1.9 Proof by contradiction

Why tell RAA from ¬I at all? Some mathematicians accept every rule of this book except RAA. For them a proof should hand over what it claims: a proof that something exists should say what exists, and a proof of "P or Q" should say which. ¬I keeps that promise; RAA can break it, and Section 8.7 shows a proof that two such numbers exist which never says which numbers they are. For now, learn to see which of the two a proof uses.

Rule	Phrases	Silent in plain prose?
RAA	"suppose, for contradiction, that $\neg P$"; "suppose not"; "assume the contrary"	never
¬¬E, LEM	"either P or not"; "P must hold, since ..."; "otherwise"	never: written as an explicit case split or an explicit removal of a double negation ("P must hold, since ...")

In analysis, RAA is the proof of Proposition 6.2.5: "suppose, for a contradiction, that $a \neq b$; then $|a-b| > 0$, so the hypothesis gives $|a-b| < |a-b|$, a contradiction; hence $a = b$". The box assumes $\neg(a = b)$ and concludes $a = b$, so it is genuine RAA; ¬I would give only $\neg\neg(a = b)$. LEM is a split such as "either the sequence has infinitely many terms in the left half of the interval, or it does not", a step of the bisection proof of the Bolzano–Weierstrass theorem. That split is genuinely non-constructive: nobody can decide in general which case holds. Section 8.7 explains why that matters, and a proof that makes such a step says so in one clause. A split such as "$x \geq 0$ or $x < 0$" is also excluded middle in logical form, but it is an instance of the trichotomy of the order of $\mathbb{R}$, one of the axioms for the real numbers, and it is never flagged.

Example. One proof by contradiction, read two ways

Let a and b be real numbers with $a \leq b + \varepsilon$ for every $\varepsilon > 0$. Prove that $a \leq b$, first by RAA; then rewrite the proof as a case split.

Solution. By RAA. Suppose, for a contradiction, that $a > b$ [assumption for RAA]. Then $\varepsilon := (a-b)/2$ is positive, so the hypothesis gives $a \leq b + (a-b)/2$ [∀E at $(a-b)/2$], that is, $2a \leq a + b$, so $a \leq b$. This contradicts $a > b$ [¬E]. Hence $a \leq b$ [RAA].

The box assumes $a > b$, which is $\neg(a \leq b)$ by the order fact of Section 2.1, and the proof concludes $a \leq b$: a negation goes **out**.

By cases. By trichotomy, $a \leq b$ or $a > b$ [∨E]. In the first case we are done. In the second, the lines of the box above reach $a \leq b$ against $a > b$, so this case cannot occur, and in it $a \leq b$ [¬E, ⊥E]. In either case $a \leq b$.

The two proofs share their mathematics, the instance $\varepsilon = (a-b)/2$, and differ only in how the classical step is written: as RAA, or as a split on trichotomy, which over $\mathbb{R}$ is an axiom. Neither is flagged: $a > b$ is $\neg(a \leq b)$ by trichotomy, so RAA adds nothing beyond that axiom. The fold below writes the same box closed by ¬I and followed by ¬¬E.

Second reading · *The same proof by ¬I and ¬¬E*

By ¬I *and* ¬¬E. Close the same box with "hence $a > b$ is impossible" [¬I], the line $\neg(a > b)$, which is $\neg\neg(a \leq b)$ under the same reading of $a > b$ as $\neg(a \leq b)$; then "so $a \leq b$ must hold" [¬¬E]. The two steps do what RAA did in one, and show that the box itself uses no classical rule.

With this third version the proofs differ only in how the classical step is written: as RAA, as ¬¬E, or as a case split. All three also read $a > b$ as $\neg(a \leq b)$ by the order fact of Section 2.1, itself an instance of trichotomy; read the other way, that fact gives $a \leq b$ from $\neg(a > b)$ directly, so the first two put an excluded-middle step on top of it. Over $\mathbb{R}$ trichotomy is an axiom, so no step here is flagged.

8.1.10 The biconditional

Rule	Phrases	Silent in plain prose?
↔I	"(⇒) … (⇐) …"; "conversely"; "we prove both directions"	never either direction
↔E	"by definition"; "that is"; "i.e."; "equivalently"	yes, for a definitional "if and only if"

In analysis, Theorem 2.4.5 of Section 2.4 (for $x, y \geq 0$, $x \leq y \iff x^2 \leq y^2$) is the model of ↔I: two labeled directions, each a →I box. ↔E is every unwinding of a definition: "f is strictly increasing, that is, $x < y$ implies $f(x) < f(y)$" (Example (A definition used in both directions)). A definition is an "if and only if" (Section 2.4), and the words "that is" use it in one direction or the other.

Example. Both directions, the second by cases

Let x be a real number with $x \geq 0$. Prove that $x \leq 1$ if and only if $x^2 \leq x$. Tag each step with its rule.

Solution. (⇒) Suppose $x \leq 1$ [assumption for →I]. Multiplying by $x \geq 0$ gives $x^2 \leq x$.

(⇐) Conversely, suppose $x^2 \leq x$ [assumption for →I]. Since $x \geq 0$, either $x = 0$ or $x > 0$ [∨E]. If $x = 0$, then $x \leq 1$. If $x > 0$, dividing $x^2 \leq x$ by x gives $x \leq 1$.

So $x \leq 1$ if and only if $x^2 \leq x$ [↔I].

The labels (⇒) and (⇐) are the two boxes of ↔I, and "conversely" opens the second.

Second reading · *A nonnegativity hypothesis, direction by direction*

The hypothesis $x \geq 0$ does a job in each: it keeps the inequality's direction when we multiply, and it reduces the cases to $x = 0$ and $x > 0$, by $\leftrightarrow$E on the definition of $\geq$. Without it, ($\Leftarrow$) still holds, since a negative x has $x^2 \geq 0 > x$, so $x^2 \leq x$ forces $x \geq 0$ and the proof above applies; but ($\Rightarrow$) fails: $x = -1$ has $x \leq 1$ and $x^2 = 1 > x$.

8.1.11 Equivalent conditions, by a cycle

Theorems often say that several conditions are equivalent ("the following are equivalent", TFAE). Proving each pair both ways would take six implications for three conditions. A **cycle** takes three.

Pattern	Phrases	Silent in plain prose?
TFAE by a cycle (Proposition 2.4.6)	"the following are equivalent"; "we prove (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii) $\Rightarrow$ (i)"	never any arrow of the cycle, the last one included

Second reading · *A cycle, worked*

Example. Three conditions, one cycle

Let n be an integer. Prove that the following are equivalent: (i) n is odd; (ii) n^2 is odd; (iii) $n^2 + 2n$ is odd. Tag the steps that carry the logic.

Solution. By Proposition 2.4.6 it suffices to prove (i) $\Rightarrow$ (ii), (ii) $\Rightarrow$ (iii) and (iii) $\Rightarrow$ (i).

- (i) $\Rightarrow$ (ii) is the direction ($\Rightarrow$) of Theorem 2.4.4.
- (ii) $\Rightarrow$ (iii): Suppose n^2 is odd [assumption for $\rightarrow$I], say $n^2 = 2a + 1$ [$\exists$E]. Then $n^2 + 2n = 2(a + n) + 1$, which is odd [$\exists$I, witness $a + n$].
- (iii) $\Rightarrow$ (i): Suppose $n^2 + 2n$ is odd. By (E1), n is even or odd [$\vee$E].
  - If n is even, say $n = 2k$ [$\exists$E], then $n^2 + 2n = 2(2k^2 + 2k)$ is even. So $n^2 + 2n$ is even and odd, which (E2) forbids [$\neg$E]: this case cannot occur [$\perp$E].
  - If n is odd, there is nothing more to show.
  - So n is odd.

The cycle closes.

Three implications did the work of six. The last arrow, (iii) $\Rightarrow$ (i), is the one a hurried writer omits. Without it the proof shows only (i) $\Rightarrow$ (ii) $\Rightarrow$ (iii), and no condition later in the chain gives back an earlier one. The case split in (iii) $\Rightarrow$ (i) uses no RAA: the impossible case reaches the goal by $\perp$E, as in Example (One proof by contradiction, read two ways)'s case version.

Quick check. To prove four conditions (i)–(iv) equivalent, a writer proves (i) ⇒ (ii), (ii) ⇒ (iii), (iii) ⇒ (iv) and (iv) ⇒ (ii). Is that enough?

Solution. No. (ii), (iii) and (iv) are shown equivalent, but nothing leads back to (i): (ii) ⇒ (i) is never shown. A fifth implication (ii) ⇒ (i) would do, or (iv) ⇒ (i) in place of (iv) ⇒ (ii).

8.1.12 The three faces of "by contradiction"

"Proof by contradiction" is used in practice for three different arguments. Each opens a box and finds $\bot$ inside it; they differ in what the box assumes and what the proof concludes. Only face (i) is always free of a classical step.

Face	The box assumes	The proof concludes	Last step	A tiny example
(i) refuting	P	$\neg P$	$\neg$I	"7 is not even"
(ii) contrapositive	$\neg Q$, inside "suppose P"	$P \to Q$, from $\neg Q \to \neg P$	Proposition 2.5.2	"if n^2 is even, then n is even"
(iii) genuine	$\neg P$	P	RAA	"if $a \le b + \varepsilon$ for every $\varepsilon > 0$, then $a \le b$"

(i) Refuting a statement. The claim is negative: irrational, "no largest", $x \neq 5$. The box assumes the positive statement, as the proof of Theorem 2.2.1 assumes $r^2 = 2$, and no classical rule is used.

(ii) The contrapositive. The claim is $P \to Q$, and the proof derives $\neg P$ from $\neg Q$, with P playing no part until the end. The passage back to $P \to Q$ is Proposition 2.5.2, which Section 2.5 derives with RAA. Written as "suppose P; suppose, for contradiction, $\neg Q$; … $\neg P$, a contradiction", it is that derivation done inline.

↻ **Second reading** · *When the contrapositive needs no classical step*

Face (ii) ends with Proposition 2.5.2, which Section 2.5 derives with RAA; that RAA is avoidable when Q is itself a negation, by the fold "RAA is never needed to finish a negation" of Section 2.5. So face (ii) uses RAA unless Q is itself a negation.

(iii) Genuine reductio. The claim P is positive, and $\neg P$ is used throughout the box, typically because it hands over an object: $a \neq b$ makes $|a - b|$ a positive number to feed to a hypothesis about every $\varepsilon > 0$.

To tell them apart, read the first sentence of the box and the sentence after it: what does it assume, and what does the proof write after it closes? The words "for contradiction" do not

decide it; a proof that $\mathbb{N}$ is not bounded above uses them for a face (i) proof: it supposes $\mathbb{N}$ bounded and concludes that it is not.

The style rule: use the weakest form that works, and say which one you use. A direct proof, when one exists, beats all three. A negative claim is proved by face (i). Face (ii) is announced as "we prove the contrapositive", which tells the reader the goal $\neg Q \to \neg P$ at once. Face (iii) is for a positive claim that resists the other routes, with its classical step made visible as in Example (One proof by contradiction, read two ways).

Quick check. Which face? (i) "Suppose $\sqrt{3} = p/q$ with integers p, q. ... a contradiction. Hence $\sqrt{3}$ is irrational." (ii) "Suppose n^2 is odd. Suppose, for a contradiction, that n is not odd. Then n is even, say $n = 2k$, so $n^2 = 4k^2 = 2(2k^2)$ is even, hence not odd, which contradicts n^2 being odd. Hence n is odd."

Solution.

(i) Face (i): the box assumes the positive statement "$\sqrt{3}$ is rational" and concludes its negation, by ¬I. (ii) Face (ii): the box assumes $\neg Q$, "n is not odd", and derives "n^2 is not odd" from it alone; the hypothesis "n^2 is odd" enters only at the clash. That is a contrapositive in costume.

8.1.13 Costumes

A proof written as "suppose, for contradiction, ..." sometimes uses its assumption for nothing except the final clash. It is then a direct proof or a contrapositive proof **in costume**, and it reads better with the costume off. Section 2.5 met the smallest case; here are the two costumes as they appear in prose, each with a test to run on a finished proof.

Example. A direct proof in costume

Here is a proof that $x^2 - 2x > 3$ for every real $x > 3$. Decide what it really uses, and rewrite it.

> Let $x > 3$. Suppose, for a contradiction, that $x^2 - 2x \leq 3$. Since $x > 3$, both $x - 3$ and $x + 1$ are positive, so $(x - 3)(x + 1) > 0$, that is, $x^2 - 2x - 3 > 0$. So $x^2 - 2x > 3$. This contradicts our assumption. Hence $x^2 - 2x > 3$.

Solution. The assumption "$x^2 - 2x \leq 3$" is cited exactly once, in "This contradicts our assumption". The sentences before it prove $x^2 - 2x > 3$ from $x > 3$ alone, so the RAA box contains a complete derivation of the goal. Deleting the box leaves a direct proof:

> Let $x > 3$. Then $x - 3 > 0$ and $x + 1 > 0$, so $x^2 - 2x - 3 = (x - 3)(x + 1) > 0$, that is, $x^2 - 2x > 3$.

Nothing is lost, the RAA is gone, and the reader no longer holds a false assumption in mind for three sentences.

The test: if the assumption of the box is cited only in the line that produces $\bot$, the box proved

the goal without it. Delete the box.

Example. A contrapositive in costume

Here is a proof that, for an integer n, if n^2 is even then n is even. Decide what it really uses, and rewrite it.

> Suppose n^2 is even. Suppose, for a contradiction, that n is not even. Then n is odd, say $n = 2k+1$ with k an integer. Then $n^2 = 4k^2+4k+1 = 2(2k^2+2k)+1$ is odd, so n^2 is not even. This contradicts the fact that n^2 is even. Hence n is even.

Solution. Here the assumption of the box, "n is not even", is used heavily; what is used only once is the **hypothesis** "n^2 is even", in "This contradicts ...". The sentences between derive "n^2 is not even" from "n is not even" alone. So the real content is $\neg Q \to \neg P$, with P the statement "n^2 is even" and Q the statement "n is even", and the frame around it is Proposition 2.5.2 done inline: face (ii). Written openly:

> We prove the contrapositive: if n is not even, then n^2 is not even. Suppose n is not even. Then n is odd, say $n = 2k+1$ with k an integer, and $n^2 = 2(2k^2+2k)+1$ is odd, so it is not even.

The rewrite tells the reader at the start what will be shown, and the final clash disappears. The classical step has not: it is now inside the announcement, which uses Proposition 2.5.2. (The proof by cases of Theorem 2.1.3 in Section 2.1 needs no RAA at all.)

The test: if everything inside the box follows from its assumption $\neg Q$ alone, and the hypothesis P of the theorem enters only in the line that produces $\bot$, then the box proved $\neg Q \to \neg P$. Present it as a contrapositive.

Second reading · *The smell the two tests share*

The two tests share one smell: *the clash at the end is with a statement the proof already had, or would have had without the box.* A genuine reductio uses its assumed negation throughout, and the clash comes somewhere the reader did not see coming; in Example (One proof by contradiction, read two ways), the assumption $a > b$ was what made $(a-b)/2$ positive.

Warning. "Assume P" means different things at the top of different proofs. At the top of a proof of $P \to Q$, it is the opening of the $\to$I box, and it is exactly right. At the top of a proof of P itself, it assumes the goal, and everything after it is circular: the last line "hence P" cites a line that is not available outside the box. For instance, "assume $\sqrt{2}$ is irrational; then ... ; hence $\sqrt{2}$ is irrational" proves nothing, however correct the middle: it is entry (N9) of the catalogue in Section 8.6.

Warning (Not allowed: connective steps in prose). Each sentence below might appear in a proof, and each is wrong. The tag names the entry of the catalogue in Section 8.6, or the item of the convention on silent steps in Section 8.3, that it breaks.

- "Suppose $x \geq 1$; then $x^3 \geq x$. Hence $x^3 \geq x$ for every real x." The line $x^3 \geq x$ belongs to a box that has closed; outside it only "if $x \geq 1$, then $x^3 \geq x$" is available: (N10). The conclusion is false at $x = \frac{1}{2}$.
- "To prove $|x| < 1$, suppose, for a contradiction, that $|x| > 1$." The negation of $|x| < 1$ is $|x| \geq 1$: (N11)(a).
- "Suppose n is odd. Then n^2 is odd.", as the whole proof of "if n^2 is even, then n is even". Without the announcement the reader sees a box that opens with neither the hypothesis nor the conclusion of the goal: the never-silent item "the announcement of a contrapositive".
- "Case 1: $x \geq 0$, so $|x| = x \geq 0$. In either case $|x| \geq 0$." There is one case and no disjunction: (N7).

8.1.14 Exercises

8.1.14.1 A. Check your understanding

Exercise. A1

(a) Name the rule that each phrase most often carries: (i) "in particular"; (ii) "conversely"; (iii) "in either case"; (iv) "this case cannot occur"; (v) "suppose, for contradiction, that x is not positive"; (vi) "we prove the contrapositive"; (vii) "that is"; (viii) "since $x \neq 0$, the first case is impossible"; (ix) "so $x \geq 0$", after "$x > 0$"; (x) "which contradicts $x^2 \geq 0$".

(b) Which of the phrases in (a) carry steps that may be silent in plain prose?

(c) A proof ends "Suppose $7 = 2k$ for an integer k. Then $k = 7/2$, which is not an integer. Hence 7 is not even." Which face of "by contradiction" is this?

(d) True or false, with a reason: a proof that begins "suppose, for contradiction" always uses RAA.

Solution. (a) (i) $\wedge$E. (ii) $\leftrightarrow$I, opening the second direction. (iii) $\vee$E, closing. (iv) $\bot$E. (v) RAA, opening its box (the goal is "x is positive", and the box assumes its negation). (vi) CP, or, as the frame of a whole proof, Proposition 2.5.2. (vii) $\leftrightarrow$E, on a definition. (viii) DS. (ix) $\vee$I. (x) $\neg$E.

(b) (i), (iii) (the closing "in either case"), (iv) (once the contradiction is stated), (vii) (for a definition), and (ix). In (viii) the name of DS may be dropped, but not the fact $x \neq 0$. The openings (ii), (v), (vi) and the clash (x) are never silent.

(c) Face (i), refuting: the box assumes the positive statement "7 is even", that is, $7 = 2k$ for some integer k, and the proof concludes its negation, by $\neg$I. (The value $7/2$ is not an integer because $3 < 7/2 < 4$.)

(d) False. The phrase also opens $\neg$I boxes. A proof that $\mathbb{N}$ is not bounded above may begin "suppose, for contradiction, that $\mathbb{N}$ is bounded above" and conclude the negation, by $\neg$I. What decides is what the box assumes and what the proof concludes.

Exercise. A2: Drill

Name the rule each phrase carries, and say whether plain prose may leave it out. Some parts ask a different question, and say so.

(a) "Suppose n is odd.", at the start of a proof of "if n is odd, then n^2 is odd".

(b) "In either case, $x^2 \geq 0$."

(c) "Since $x \neq 0$, the first case is impossible, so $x = 3$", after "$x = 0$ or $x = 3$".

(d) "$x > 2$, so certainly $x \geq 2$."

(e) "This contradicts $n^2 \geq 0$."

(f) "Conversely, suppose $x^2 = 4$ and $x > 0$."

(g) "Suppose, for a contradiction, that n is not prime", in a proof that ends "Hence n is prime."

(h) A proof that $\sqrt{5}$ is irrational opens: "Suppose, for a contradiction, that $\sqrt{5}$ is rational." Which annotation would its derivation print beside that assumption?

(i) "Suppose 15 is a perfect square. ... a contradiction. Hence 15 is not a perfect square." Which face of "by contradiction" is this?

(j) A proof of "if $x^2 < 1$, then $x > -1$" begins "Suppose $x \leq -1$. Then $-x \geq 1$, so $x^2 = (-x)^2 \geq 1$." What should come before it?

(k) "Let n be even, say $n = 2k$. Suppose, for a contradiction, that n^2 is not even. Then $n^2 = 4k^2 = 2(2k^2)$ is even, which contradicts n^2 not being even. Hence n^2 is even." Is this proof in costume?

(l) Four conditions are to be shown equivalent. How many implications does a cycle need, and how many would proofs of every ordered pair need?

Solution.

(a) $\to$I, opening; never silent.

(b) $\vee$E, closing; may be silent.

(c) DS; its name may go, the fact $x \neq 0$ may not.

(d) $\vee$I ($x \geq 2$ is $x > 2 \vee x = 2$); may be silent.

(e) $\neg$E; never silent: both lines of the clash are named.

(f) $\leftrightarrow$I, opening the second direction; never silent.

(g) RAA, opening; never silent.

(h) "asm, for $\neg$I". The goal, "$\sqrt{5}$ is irrational", is the negation "$\sqrt{5}$ is not rational". The box assumes the positive statement and is closed by $\neg$I, whatever the words "for a contradiction" suggest: Section 2.5's distinction, face (i) of "The three faces". (No fact about $\sqrt{5}$ is needed.)

(i) Face (i), refuting: $\neg$I.

(j) The announcement "We prove the contrapositive" (CP; never silent).

(k) Yes, a direct proof in costume: the assumption is cited only at the clash; delete the box. (The box opens with the negation of the goal, so the item is not an instance of (N11)(a).)

(l) Four for a cycle. Twelve for every ordered pair ($4 \cdot 3$), that is, six biconditionals.

8.1.14.2 B. Practice

Exercise. B1: Take off the costume

Here is a proof that $\frac{a+b}{2} \geq \sqrt{ab}$ for all real $a > 0$ and $b > 0$.

> Let $a, b > 0$. Suppose, for a contradiction, that $\frac{a+b}{2} < \sqrt{ab}$. Since $(\sqrt{a}-\sqrt{b})^2 \geq 0$, expanding gives $a-2\sqrt{ab}+b \geq 0$, so $\frac{a+b}{2} \geq \sqrt{ab}$. This contradicts our assumption. Hence $\frac{a+b}{2} \geq \sqrt{ab}$.

(a) Which costume is this? Justify your answer by saying where the assumption of the box is cited.

(b) Rewrite the proof without the contradiction.

Solution. (a) A direct proof in costume (Example (A direct proof in costume)). The assumption $\frac{a+b}{2} < \sqrt{ab}$ is cited only in "This contradicts our assumption", the line that produces $\bot$. The sentence beginning "Since" proves the goal from $a, b > 0$ alone, without using the assumption.

(b) Let $a, b > 0$, so that $\sqrt{a}$ and $\sqrt{b}$ are defined and $\sqrt{a}\sqrt{b} = \sqrt{ab}$. Since $(\sqrt{a} - \sqrt{b})^2 \geq 0$, expanding gives $a - 2\sqrt{ab} + b \geq 0$. Adding $2\sqrt{ab}$ to both sides and dividing by 2 gives $\frac{a+b}{2} \geq \sqrt{ab}$, as claimed.

Exercise. B2: Where the classical step sits

Here is a proof that if x is a real number with $|x| \leq |y|$ for every real $y \neq 0$, then $x = 0$.

> Suppose $|x| \leq |y|$ for every real $y \neq 0$. Suppose, for a contradiction, that $x \neq 0$. Then $y := x/2$ is a non-zero real number, so the hypothesis gives $|x| \leq |x/2| = |x|/2$, and hence $|x| \leq 0$. But $|x| > 0$, since $x \neq 0$. This is a contradiction. Hence $x = 0$.

(a) Is the last step RAA or ¬I? Is the proof in costume?

(b) Rewrite it so that the box closes by ¬I and one further step removes a double negation. Tag the two steps.

(c) Which of the two versions tells the reader more, and what?

Solution. (a) RAA: the box assumes $\neg(x = 0)$, written $x \neq 0$, and the proof concludes the positive statement $x = 0$. It is not in costume. The assumption does work inside the box (it makes $x/2$ non-zero, so that the hypothesis applies to it), and the hypothesis is used in the middle of the box, not only at the clash; neither test of this section is met.

(b) Suppose $|x| \leq |y|$ for every real $y \neq 0$. If $x \neq 0$, then $x/2 \neq 0$, so the hypothesis gives $|x| \leq |x|/2$, that is, $|x| \leq 0$, against $|x| > 0$. Hence $x \neq 0$ is impossible [¬I], so $x = 0$ must hold [¬¬E].

(c) The second. It shows that the box uses no classical rule: without one, the argument proves $\neg\neg(x = 0)$, "x cannot be non-zero", and the classical rule enters only in the final step, "so $x = 0$ must hold" [¬¬E]. The first version does both steps at once in the single RAA,

and so hides where the classical rule enters.

Exercise. B3: Two clauses of a least upper bound

Let $A \coloneqq \{x \in \mathbb{R} : 0 < x < 2\}$. Prove that 2 is an upper bound of A, and that no real number $t < 2$ is an upper bound of A. Tag each step with its rule.

Hint: in the second part, take a number halfway between $\max\{t, 0\}$ and 2.

Solution. *(a)* Let $a \in A$ [$\forall$I]. Then $0 < a < 2$, by the definition of A; in particular $a < 2$ [$\wedge$E], so $a \leq 2$ [$\vee$I]. So 2 is an upper bound of A.

(b) Let t be a real number with $t < 2$ [$\forall$I, $\to$I], and suppose t is an upper bound of A [assumption for $\neg$I]. Let $m \coloneqq \max\{t, 0\}$ and $a \coloneqq (m + 2)/2$. We have $t \geq 0$ or $t < 0$, and by the definition of the maximum we consider the two cases [$\vee$E]: if $t \geq 0$, then $m = t$, so $m \geq t$, $m \geq 0$ and $m < 2$; if $t < 0$, then $m = 0$, so $m > t$, $m \geq 0$ and $m < 2$. So in either case $m \geq t$ and $0 \leq m < 2$. Hence $1 \leq a < 2$; thus $0 < a$ and $a < 2$, that is, $a \in A$ [$\wedge$I]. As t is an upper bound of A, $a \leq t$ [$\forall$E at a]. But $a - m = (2 - m)/2 > 0$, so $a > m \geq t$, and $a > t$. This contradicts $a \leq t$ [$\neg$E]. Hence t is not an upper bound of A [$\neg$I].

So 2 is an upper bound of A, and no smaller number is [$\wedge$I].

The $\wedge$E of (a) could have been silent ("then $a < 2$, so $a \leq 2$"); the $\vee$I after it weakened a strict inequality, since $a \leq 2$ means $a < 2 \vee a = 2$. In (b), the two conjuncts $0 < a$ and $a < 2$ were displayed and then merged by "that is, $a \in A$". The bound $m < 2$ needed the case split hidden in max, and it is written out, as the warning on "clearly" in "Disjunction and cases" requires. The final sentence is the $\wedge$I of the two clauses. And the clashing lines $a \leq t$ and $a > t$ are both named before "contradicts", as the $\neg$E row of the dictionary requires.

Exercise. B4: Two integers of the same parity

Let a and b be integers. Prove that the following are equivalent, by a cycle: (i) $a + b$ is even; (ii) $a - b$ is even; (iii) a and b are both even or both odd. Say which arrow closes the cycle, and tag the logical steps of that arrow.

Solution. By Proposition 2.4.6, three implications suffice.

- (i) $\Rightarrow$ (ii): Suppose $a + b$ is even; choose k with $a + b = 2k$ [$\exists$E]. Then $a - b = (a + b) - 2b = 2(k - b)$, so $a - b$ is even, with the witness $k - b$ [$\exists$I].
- (ii) $\Rightarrow$ (iii): Suppose $a - b$ is even, and choose j with $a - b = 2j$ [$\exists$E]. By (E1), b is even or odd [$\vee$E].
  - If $b = 2m$ [$\exists$E], then $a = 2(j + m)$ is even [$\exists$I], so both are even [$\wedge$I, then $\vee$I].
  - If $b = 2m + 1$ [$\exists$E], then $a = 2(j + m) + 1$ is odd [$\exists$I], so both are odd [$\wedge$I, then $\vee$I].
- (iii) $\Rightarrow$ (i), the closing arrow: Suppose a and b are both even or both odd [assumption for $\to$I]. There are two cases [$\vee$E].
  - Both even: choose s and t with $a = 2s$ and $b = 2t$ [$\exists$E twice, two fresh names]. Then $a + b = 2(s + t)$ [$\exists$I, witness $s + t$].

- Both odd: choose s and t with $a = 2s + 1$ and $b = 2t + 1$ [∃E twice]. Then $a + b = 2(s + t + 1)$ [∃I, witness $s + t + 1$].
- In either case $a + b$ is even [∨E, closing; may be silent in plain prose].

The closing arrow is (iii) ⇒ (i). Its logic is a two-case ∨E, each case with two ∃E boxes under distinct names and one ∃I; no RAA is used. (The names s, t may be reused in the second case: sibling boxes may share a name, as (N4) of Section 8.6 notes.)

Marking scheme. (10 points) Points are given per role, and a correct proof by another cycle earns the same points for the corresponding arrows. - 2: an arrow proved by one ∃E and one ∃I (here (i) ⇒ (ii), with k chosen and the witness $k - b$). - 3: an arrow from a parity hypothesis, by a split with (E1) named (1) and a witness in each case (2) (here (ii) ⇒ (iii), witness $j + m$). - 3: an arrow from (iii): the two cases (1); two distinct names in each case (1); the witnesses (1). - 1: the cycle named, with Proposition 2.4.6 or the sentence that three implications suffice, and the closing arrow identified. - 1: the last sentence says what was shown.

Fatal errors (0 for the step): only two arrows proved; one name for both witnesses, "$a = 2s$ and $b = 2s$" (N4); "$a - b$ is even, so a and b have the same parity" given as its own reason, with neither proof nor citation (N9).

8.1.14.3 C. Going deeper

Exercise. C1: Euclid without the contradiction

Here is a common proof that there are infinitely many primes. It uses two facts about integers that we take as known here: every integer $m \geq 2$ has a prime divisor (*Sets, Relations, and Cardinality*, Chapter 2 proves it); and if an integer d divides two integers, it divides their difference (if $u = dk$ and $v = dl$, then $u - v = d(k - l)$).

> Suppose, for a contradiction, that there are only finitely many primes, say $p_1, \ldots, p_k$. Let $N := p_1 p_2 \cdots p_k + 1$. Then $N \geq 2$, so N has a prime divisor p. Since $p_1, \ldots, p_k$ are all the primes, $p = p_i$ for some i. Then p_i divides both N and $p_1 p_2 \cdots p_k$, so it divides their difference 1, which is impossible for a prime. This is a contradiction. Hence there are infinitely many primes.

(a) "There are infinitely many primes" means "the set of primes is **not** finite". Is the last step of the proof RAA or ¬I?

(b) Where is the assumption "$p_1, \ldots, p_k$ are all the primes" used? Which smell from this section does that suggest?

(c) Rewrite the argument as a direct proof of: *for every finite list $p_1, \ldots, p_k$ of primes, with $k \geq 1$, there is a prime not in the list.*

(d) Say what the rewrite gains over the original.

Solution. (a) ¬I. The box opens with "the set of primes is finite", the positive statement, and the conclusion is its negation: face (i). No RAA is used.

(b) Once, in "$p = p_i$ for some i", where the clash is manufactured. Everything else holds for **any** finite list of primes, complete or not: the smell of a costume.

(c) Let $p_1, \ldots, p_k$ be a finite list of primes, $k \geq 1$, and let $N := p_1 p_2 \cdots p_k + 1$. Each $p_i \geq 2$,

so $N \geq 3$, and N has a prime divisor p. Let i be an index with $1 \leq i \leq k$ [$\forall$I, opening], and suppose $p = p_i$ [assumption for $\neg$I]. Then p divides N and $p_1 p_2 \cdots p_k$, so it divides their difference 1, say $1 = pm$ with m an integer. Since $p > 0$, $m > 0$, so $m \geq 1$ and $1 = pm \geq p \geq 2$, a contradiction [$\neg$E]. Hence $p \neq p_i$ [$\neg$I]. Since i was arbitrary, $p \neq p_i$ for every i [$\forall$I, closing], and p is a prime not in the list.

(d) The rewrite proves a stronger statement: given any finite list of primes, it **produces** a prime outside it, any prime divisor of N; the original shows only that a list of all the primes is impossible. Its one $\neg$I box, for the negative claim $p \neq p_i$, is honest, since its assumption is what the clash uses. The original theorem follows at once: a finite list of all the primes is non-empty, since 2 is prime, so by (c) it would miss a prime.

Marking scheme. (10 points) - 2: (a) $\neg$I, with the reason (the box assumes the positive statement and the proof concludes its negation). - 2: (b) locates the single use of the assumption and names the costume smell. - 4: (c) a proof for an arbitrary finite list: N defined, $N \geq 2$ justified, a prime divisor p taken, $p \neq p_i$ for every i shown. - 2: (d) names the gain: a prime outside any given list is produced. Fatal errors (0 for the step): in (c), a proof that still assumes the list contains all the primes; in (a), calling the step RAA because of the words "for a contradiction".

Exercise. C2: Which face, and the weakest form

Here is a proof that if x is a real number with $x^2 < \varepsilon$ for every $\varepsilon > 0$, then $x = 0$.

> Suppose $x^2 < \varepsilon$ for every $\varepsilon > 0$. Suppose, for a contradiction, that $x \neq 0$. Then $x^2 > 0$, so we may take $\varepsilon = x^2$, and the hypothesis gives $x^2 < x^2$, which is impossible. Hence $x = 0$.

(a) Which face of "by contradiction" is this, as written?

(b) Rewrite the proof in the weakest form you can, and say which rules it uses at its logical steps, and which school fact.

Solution (Hint 1). The goal $x = 0$ is positive. Look for a school fact about x^2 that is a disjunction: one case leads to the goal through another school fact, and the hypothesis rules out the other.

Solution (Hint 2). Start: "Since $x^2 \geq 0$, either $x^2 > 0$ or $x^2 = 0$."

Solution. (a) Face (iii), genuine reductio. The box assumes $\neg(x = 0)$, written $x \neq 0$, uses it to make x^2 a positive number to which the hypothesis applies, and concludes the positive statement $x = 0$. The last step is RAA. It is not in costume: the assumption does work inside the box.

(b) Suppose $x^2 < \varepsilon$ for every $\varepsilon > 0$ [assumption for $\to$I]. Since $x^2 \geq 0$, either $x^2 > 0$ or $x^2 = 0$, and we consider the two cases [$\vee$E].

Case $x^2 > 0$. Then the hypothesis applies to $\varepsilon = x^2$ [$\forall$E] and gives $x^2 < x^2$, which contradicts

the fact that no number is less than itself [¬E]. So this case cannot occur, and in it $x = 0$ [⊥E].

Case $x^2 = 0$. Then $x \cdot x = 0$, so $x = 0$ [→E], since a product of two real numbers is 0 only if a factor is.

In either case $x = 0$ [∨E].

The logical steps are →I, ∨E, ∀E, ¬E, →E (citing the second school fact, that a product is 0 only if a factor is) and ⊥E; no classical rule is used. The classical content has moved into the school fact "$x^2 > 0$ or $x^2 = 0$", which rests on the trichotomy of the order, one of the axioms for the real numbers; Section 8.7 explains why that axiom is where the book makes its classical commitment.

Marking scheme. (10 points) - 3: (a) identifies RAA (face (iii)), with the reason: the assumption is a negation and the conclusion is positive. - 2: (b) splits on a disjunction that is justified (from $x^2 \geq 0$, or from trichotomy). - 3: (b) closes the case $x^2 > 0$ by the instance $\varepsilon = x^2$, a named clash, and ⊥E. - 2: (b) settles the case $x^2 = 0$ with its reason and concludes "in either case". A rewrite with ¬I and ¬¬E in place of the case split earns the points of (b) except the 2 for the split, provided it names ¬¬E as its one classical step. Fatal errors (0 for the step): calling (a) face (i) because the word "contradiction" appears; splitting on "$x^2 > 0$ or $x^2 < 0$", which does not exhaust.

8.2 The Quantifier Rules in Prose

Section 8.1 read the connective rules in the sentences of written proofs. The statements of analysis, though, are made of "for every" and "there exists", and most of the words in an analysis proof carry a quantifier rule: "let", "fix", "choose", "take", "will do". This section completes the dictionary with the rows for the quantifiers, equality and unique existence, for definitions, "the" and double inclusion, and for the patterns that are not single rules but are written as if they were: the lemma ("it suffices to show") and its two "stronger" cases, induction, and symmetry ("without loss of generality"). It also makes exact one habit of analysis prose, the parenthesis "(depending on ε)".

8.2.1 Three sentences, three rules

Section 6.1 derived "if $a_n \to L$, then $2a_n \to 2L$" in sixteen lines. Written the way a mathematician writes it, the same argument takes three sentences:

> Let $\varepsilon > 0$. Since $a_n \to L$ and $\varepsilon/2 > 0$, we may choose $N \in \mathbb{N}$ (depending on ε) such that $|a_n - L| < \varepsilon/2$ for every $n \geq N$. Then for every $n \geq N$, $|2a_n - 2L| = 2|a_n - L| < \varepsilon$.

Here is the derivation of Section 6.1 again, with the same line numbers. To keep its lines short, write $A(n, \varepsilon)$ for $|a_n - L| < \varepsilon$ and $B(n, \varepsilon)$ for $|2a_n - 2L| < \varepsilon$.

1		$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to A(n, \varepsilon))$	
2	ε_0	$\varepsilon_0 > 0$	asm, for $\forall$I
3		$\varepsilon_0/2 > 0$	arith 2
4		$\exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to A(n, \varepsilon_0/2))$	$\forall$E 1, 3
5		N_0 $N_0 \in \mathbb{N}$	asm, for $\exists$E
6		$\forall n \in \mathbb{N}\ (n \geq N_0 \to A(n, \varepsilon_0/2))$	asm, for $\exists$E
7		n_0 $n_0 \in \mathbb{N}$	asm, for $\forall$I
8		$n_0 \geq N_0 \to A(n_0, \varepsilon_0/2)$	$\forall$E 6, 7
9		$n_0 \geq N_0$	asm, for $\to$I
10		$A(n_0, \varepsilon_0/2)$	$\to$E 8, 9
11		$B(n_0, \varepsilon_0)$	arith 10
12		$n_0 \geq N_0 \to B(n_0, \varepsilon_0)$	$\to$I 9–11
13		$\forall n \in \mathbb{N}\ (n \geq N_0 \to B(n, \varepsilon_0))$	$\forall$I 7–12
14		$\exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to B(n, \varepsilon_0))$	$\exists$I 5, 13
15		$\exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to B(n, \varepsilon_0))$	$\exists$E 4, 5–14
16		$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to B(n, \varepsilon))$	$\forall$I 2–15

Sentence	Lines
"Let $\varepsilon > 0$."	opens the box 2–15; its closing line 16 is not written
"Since $a_n \to L$ and $\varepsilon/2 > 0$, we may choose N . . . "	3 and 4, then opens the box 5–14
"Then for every $n \geq N$, $\lvert 2a_n - 2L\rvert = 2\lvert a_n - L\rvert < \varepsilon$."	7–13
(the paragraph ends)	14, 15, 16

Each sentence opens a box, and the boxes are nested in the order of the sentences.

- "Let $\varepsilon > 0$" opens the bounded $\forall$I box of lines 2–15 in the derivation above, with the fresh name ε_0. Its closing line, line 16, is not written: the paragraph simply ends, and a reader who heard "let" knows that the box closes there.
- "Choose N . . . " is two rules. The words "since $a_n \to L$ and $\varepsilon/2 > 0$" are the $\forall$E of line 4, at the term $\varepsilon_0/2$, with its bound from line 3. The word "choose" opens the $\exists$E box of lines 5–14, with the fresh name N_0.
- "Then for every $n \geq N$" opens the $\forall$I box of n_0 and the $\to$I box of $n_0 \geq N_0$ (lines 7 and 9), and the computation is lines 8, 10 and 11. The words "for every $n \geq N$" also announce the lines that close these two boxes, 12 and 13.
- The $\exists$I of line 14, which writes "some N works" from "this N works", and the closing line 15 of the $\exists$E box, are silent. They are the end of the paragraph.

The one word in the paragraph that no rule asks for is the parenthesis "(depending on ε)". It is the subject of a subsection below. First, the rows.

8.2.2 The quantifier rows

Rule	Phrases	Silent in plain prose?
$\forall$I	"let $x \in A$"; "fix $\varepsilon > 0$"; "let $\varepsilon > 0$ be given"; "since x was arbitrary"	never the introduction; "since x was arbitrary" may be silent
$\forall$E	"in particular, for $x = \dots$"; "applying … with $\varepsilon/2$ in place of ε"	the instance must be named when it is not the obvious one
$\exists$I	"take $N = \dots$"; "let $\delta := \dots$; then …"; "… will do"	never; its dependencies are stated
$\exists$E	"choose N such that"; "by hypothesis there is k with $n = 2k$"; "let x_0 be such a point"	never; the name must be fresh

The phrases in the table are mostly the bounded forms of the rules (Proposition 6.1.1): "let $x \in A$" opens a $\forall$I box whose assumption is $x \in A$, and "choose $N \in \mathbb{N}$ such that …" opens an $\exists$E box with the two assumptions $N \in \mathbb{N}$ and "…". A bound is part of the phrase: "let $\varepsilon > 0$" is one step, not "let ε" followed by "suppose $\varepsilon > 0$".

In analysis, $\forall$I is the first sentence of every limit proof: "Let $\varepsilon > 0$." $\forall$E is "apply the definition of $a_n \to L$ with $\varepsilon/2$": the instance $\varepsilon/2$ is not the obvious one, so it is named. $\exists$I is the line that places a constant, "let $\delta := \min\{1, \varepsilon/10\}$", written before the computation that shows it works. $\exists$E is "choose N (depending on ε) such that …", in which the name N has not been used before in the proof.

Example. A constant sequence

Let c be a real number and let $a_n := c$ for every $n \in \mathbb{N}$. Prove that $a_n \to c$, and annotate each phrase with its rule.

Solution. Let $\varepsilon > 0$ [$\forall$I, opening]. Take $N := 0$ [$\exists$I]. Let $n \in \mathbb{N}$ with $n \geq N$ [$\forall$I, $\to$I, opening]. Then

$$|a_n - c| = |c - c| = 0 < \varepsilon,$$

where the first equality substitutes $a_n = c$ [=E] and the last step uses $\varepsilon > 0$, the assumption of the first box. This proves $a_n \to c$.

The last sentence stands for three closing lines, all silent: the $\to$I and $\forall$I for n ("for every $n \geq N$"), and the $\forall$I for ε ("since ε was arbitrary"). The witness "take $N := 0$" is placed before the work that justifies it.

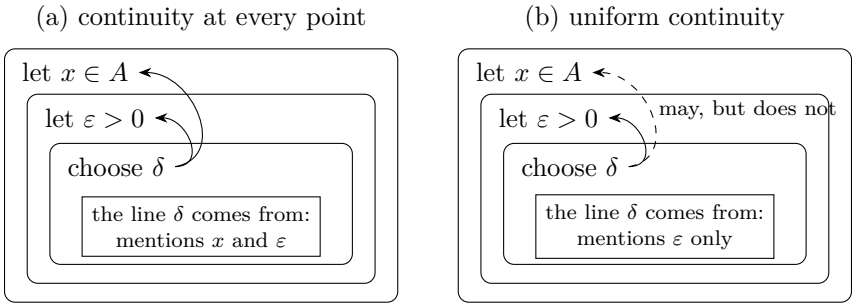
Two things here are degenerate, and both are worth seeing once. The witness N depends on nothing: any $N \in \mathbb{N}$ works. And the assumption $n \geq N$ is never used. The box is still opened, because the goal $\forall n \in \mathbb{N}\ (n \geq N \to \dots)$ asks for it; without "let $n \geq N$", the line "$|a_n - c| = 0$" would not say which n it is about.

8.2.3 The dependency parenthesis

In the paragraph that opened this section, N was chosen "(depending on ε)". Section 6.4 read off a derivation the rule behind the parenthesis:

A constant chosen in a box may depend on the names of the boxes open around that box, and it does depend only on the names in the line it comes from.

In the derivation of Section 6.1, the witness N_0 is obtained at line 5 from line 4, which mentions ε_0, inside the box of ε_0 only. So N_0 may depend on ε_0, and it does. In the derivation of Example (Uniformly continuous implies continuous), the witness δ_0 is obtained inside the boxes of x_0 and ε_0, so it **may** depend on both; but it comes from a line that mentions ε_0 alone, so it **does** depend only on ε_0, and that is uniformity.



*A chosen constant **may** depend on the names of the boxes it sits in, and **does** depend only on the names in the line it comes from (the solid arrows).*

Numbers show why the arrows matter. Take $f(x) = x^2$ and $\varepsilon = 1$. At the point $a \geq 0$, the largest δ with $|x^2 - a^2| < 1$ for every x with $|x - a| < \delta$ is $\sqrt{a^2+1} - a$: about 0.41 at $a = 1$, and about 0.005 at $a = 100$. No single δ serves every point, so δ must depend on the point as well as on ε: panel (a). For $f(x) = 3x + 1$, $|f(x) - f(a)| = 3|x - a|$, so $\delta = \varepsilon/3$ works at every point at once: δ depends on ε only, panel (b).

The parenthesis is the prose record of this. "Choose N (depending on ε)" says: the boxes open here are those of ε (besides the data of the theorem, such as the sequence), and the line from which N comes mentions ε. "Choose δ (depending on ε and on x)" says that δ came from a line mentioning both, inside both boxes. A witness that came from a line mentioning neither is a constant, and the proof may say "(depending on neither)" or simply "a fixed δ". This is why proofs in analysis say what each chosen constant depends on. A missing parenthesis is not a wrong step, but it hides the one fact a reader most needs in order to check the quantifier order, and many false proofs in analysis go wrong exactly there.

Quick check. For $f(x) = x^2$ and $\varepsilon = 1$, does $\delta = 0.1$ work at the point $a = 2$? At $a = 10$?

Solution. At $a = 2$, yes, since $0.1 < \sqrt{5} - 2 \approx 0.236$. At $a = 10$, no: $x = 10.06$ has $|x - 10| = 0.06 < 0.1$ and $x^2 - 100 = 1.2036 \geq 1$.

Example. The reciprocals tend to zero

Prove that $1/n \to 0$, for the sequence $(1/n)_{n\geq 1}$. Annotate each phrase with its rule, and

state what the chosen constant depends on.

Solution. The sequence starts at $n = 1$, so in the definition of Section 6.1 the numbers N and n range over the positive integers $\mathbb{Z}_{>0}$. We use one school fact about the real numbers: for every real number u there is an integer greater than u. (It follows from the axioms for the real numbers; here we take it as known.)

Let $\varepsilon > 0$ [$\forall$I, opening]. It suffices to find $N \in \mathbb{Z}_{>0}$ with

$$\frac{1}{N} < \varepsilon, \tag{$*$}$$

since then every $n \in \mathbb{Z}_{>0}$ with $n \geq N$ has $0 < 1/n \leq 1/N < \varepsilon$, so $|1/n - 0| < \varepsilon$ [it suffices: the claim $(*)$ displayed, then used].

By the school fact, applied to $u = 1/\varepsilon$ [$\forall$E], there is an integer greater than $1/\varepsilon$; choose such an integer N [$\exists$E], which depends on ε. Since $\varepsilon > 0$, we have $1/\varepsilon > 0$, so $N > 0$ and $N \in \mathbb{Z}_{>0}$. From $N > 1/\varepsilon > 0$, taking reciprocals reverses the inequality: $1/N < \varepsilon$, which is $(*)$. So this N will do [$\exists$I], and $1/n \to 0$.

Read the proof as a skeleton. The first sentence opens the box of ε. The second is an "it suffices to show", in the sense of the row below: it states a simpler claim $(*)$, shows on the spot that $(*)$ is enough, and leaves $(*)$ to be proved. The next paragraph proves $(*)$. It contains two different quantifier steps that both concern N: "choose" is the $\exists$E on the school fact, which hands over some integer beyond $1/\varepsilon$ and calls it N; "will do" is the $\exists$I that offers this N as the witness the definition asks for. Between them, the clause "so $N > 0$ and $N \in \mathbb{Z}_{>0}$" checks the bound $N \in \mathbb{Z}_{>0}$ of the witness, which the bounded $\exists$I needs as a line.

The constant N depends on ε only: it came from the school fact at $1/\varepsilon$, inside the box of ε. For instance, $\varepsilon = 0.03$ has $1/\varepsilon = 33.3\ldots$, and $N = 34$ is a choice that works; so is any larger integer.

Quick check. "For every $\varepsilon > 0$, $|a_n - L| < \varepsilon$ for all large n." Where is the hidden $\exists N$, and between which two quantifiers does it stand?

Solution. "For all large n" is "there is an N such that for every $n \geq N$" (Section 6.1): the hidden $\exists N$ stands after $\forall \varepsilon > 0$ and before $\forall n \geq N$. Written out, the sentence is $\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to |a_n - L| < \varepsilon)$, the definition of $a_n \to L$. Since $\exists N$ comes after $\forall \varepsilon$, the N may depend on ε.

8.2.4 Equality and unique existence

Rule	Phrases	Silent in plain prose?
=I, =E	"substituting"; chains "$\cdots = \cdots = \ldots$"; "by (1)"	yes, inside chains

Rule	Phrases	Silent in plain prose?
$\exists!$	"there is exactly one"; "existence … uniqueness …"; "suppose y also …"	never either half

In analysis, $=$E is at work in every chain of equalities and inequalities, such as

$$|x^2 - 4| = |x-2|\,|x+2| \le 5|x-2| \quad \text{for } |x-2| < 1,$$

where each $=$ link substitutes equals for equals and each inequality link applies one fact about order (Section 5.7). (Here $|x-2| < 1$ gives $1 < x < 3$, so $3 < x+2 < 5$.) The $=$I lines $t = t$ that such chains rest on are never written. Unique existence (Definition 5.8.1) has two halves, which Proposition 5.8.1 separates: "$2x + 3 = 7$ has exactly one real solution: $x = 2$ is one, and if $2y + 3 = 7$, then $y = 2$". The first clause is existence, by $\exists$I; the second is uniqueness, a $\forall$I and $\to$I box on a second candidate. Neither half implies the other, so neither may be silent.

The uniqueness half often stands alone, when existence is not claimed or is proved elsewhere. Uniqueness of limits is a case of this kind.

Example. A sequence has at most one limit

Let (a_n) be a sequence of real numbers with $a_n \to L$ and $a_n \to M$. Prove that $L = M$, and annotate each phrase with its rule.

Solution. Let $\varepsilon > 0$ [$\forall$I, opening]. Then $\varepsilon/2 > 0$, so, applying the definition of $a_n \to L$ with $\varepsilon/2$ in place of ε [$\forall$E], we may choose $N_1 \in \mathbb{N}$ (depending on ε) such that $|a_n - L| < \varepsilon/2$ for every $n \ge N_1$ [$\exists$E, fresh name N_1]. Likewise, from $a_n \to M$, choose $N_2 \in \mathbb{N}$ (depending on ε) such that $|a_n - M| < \varepsilon/2$ for every $n \ge N_2$ [$\forall$E, $\exists$E, fresh name N_2]. Let $n := \max\{N_1, N_2\}$. Then $n \in \mathbb{N}$, $n \ge N_1$ and $n \ge N_2$, so both bounds apply at n [$\forall$E at n, twice, with $\to$E], and

$$\begin{aligned} |L - M| &= |(L - a_n) + (a_n - M)| \\ &\le |a_n - L| + |a_n - M| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

where the inequality is the triangle inequality of Example (The triangle inequality by cases), with $|L - a_n| = |a_n - L|$. Since $\varepsilon > 0$ was arbitrary, $|L - M| < \varepsilon$ for every $\varepsilon > 0$ [$\forall$I, closing], and Proposition 6.2.5 gives $L = M$ [$\to$E, citing it].

Four things in this proof are never silent, and each is visible. The instance $\varepsilon/2$ is named twice, because it is not the obvious instance. The two witnesses get two names, N_1 and N_2: the hypotheses promise a threshold for L and a threshold for M, and nothing says they are the same number, so calling both N would assert an equality nobody proved (Section 8.6 catalogues the error). The point $n = \max\{N_1, N_2\}$ is named, since the bounds are applied at it. And the closing result is cited by name.

↻ **Second reading** · *The bookkeeping of the two choices*

What is silent is the bookkeeping of the two $\exists$E boxes. The line $|L - M| < \varepsilon$ mentions neither N_1, N_2 nor n, so the side condition of $\exists$E lets it leave both boxes, and then the box of ε closes on

it. The calculation chain hides its =E steps, as the row allows. The proof shows that a sequence has **at most one** limit: the uniqueness half of ∃!. Whether a given sequence has a limit at all is a separate question, which this proof does not touch.

8.2.5 Definitions, descriptions and sets

Three more phrases do logical work in almost every chapter of analysis: "we define", "the", and "the two sets are equal".

Pattern	Phrases	Silent in plain prose?
a definition (a definitional extension, Proposition 5.8.2)	"we define $f(x) \coloneqq \dots$"; "call n *perfect* if $\dots$"; "write d for $\dots$"	never the definition; its uses ($\leftrightarrow$E or =E with it) yes
"the" (a definite description, Definition 5.8.3)	"let L be the limit of (a_n)"; "the least element of A"; "$\lim a_n$"	never the existence and uniqueness behind "the": proved or cited before the name is used
double inclusion	"we show $A \subseteq B$ and $B \subseteq A$"; element chasing: "let $x \in A$; $\dots$ so $x \in B$"; "so $A = B$"	never either inclusion, nor the opening "let $x \in \dots$" of each, nor the appeal to extensionality that turns "the same elements" into $A = B$

"The", as in "the limit", names an object. It may be used only once there is exactly one such object (Section 5.8): for a limit, existence is the hypothesis that (a_n) converges, and uniqueness is Example (A sequence has at most one limit). Chapter 5 read $A \subseteq B$ as $\forall x\,(x \in A \to x \in B)$. Two inclusions give $\forall x\,(x \in A \leftrightarrow x \in B)$, by $\leftrightarrow$I inside $\forall$I: the two sets have the same elements. Each inclusion is an **element chase**: "let $x \in A$" opens a bounded $\forall$I box whose last line is $x \in B$. The last step, from "the same elements" to $A = B$, is not logic. Logic gives only the converse, that equal sets have the same elements (Example (Equal Sets Have the Same Elements)). The step itself is the **axiom of extensionality**, an axiom about sets (Section 5.7), and a proof cites it, as a premise or with the words "by extensionality"; it is never silent.

Second reading · *What a definition adds*

A definition adds a word and proves nothing new about the old words (Proposition 5.8.2, Section 5.8). So "we define" writes an abbreviation, and every later use of it is $\leftrightarrow$E or =E with the definition.

Example. Two inclusions

Let $A \coloneqq \{x \in \mathbb{R} : x^2 \leq 2x\}$ and $B \coloneqq \{x \in \mathbb{R} : 0 \leq x \leq 2\}$. Prove that $A = B$, and tag the steps that carry the logic.

Solution.

- ($\subseteq$) Let $x \in A$ [assumption for $\forall$I]; so $x^2 \leq 2x$, that is, $x(x-2) \leq 0$.
  - If $x < 0$, then $x - 2 < 0$ too, and a product of two negative numbers is positive: $x(x-2) > 0$. This clashes with $x(x-2) \leq 0$ [$\neg$E], so $x < 0$ is impossible [$\neg$I], and $x \geq 0$ by trichotomy.
  - If $x > 2$, then both factors are positive and the same clash follows. So $x \leq 2$.
  - Hence $0 \leq x \leq 2$ [$\wedge$I], and $x \in B$.
- ($\supseteq$) Let $x \in B$ [assumption for $\forall$I]; so $x \geq 0$ and $x - 2 \leq 0$ [$\wedge$E]. A product of a non-negative and a non-positive number is at most 0, so $x(x-2) \leq 0$, that is, $x^2 \leq 2x$, and $x \in A$.
- So $A \subseteq B$ and $B \subseteq A$: every real x has $x \in A \leftrightarrow x \in B$ [$\leftrightarrow$I inside $\forall$I]. By extensionality, $A = B$ [$\forall$E, $\rightarrow$E with the axiom].

The words "that is" unwind the two sets' definitions ($\leftrightarrow$E, silent at level 3).

Quick check. To prove $\{n \in \mathbb{Z} : n^2 \text{ is even}\} = \{n \in \mathbb{Z} : n \text{ is even}\}$ by double inclusion, which inclusion is a theorem of Chapter 2, and which takes one line?

Solution. $\subseteq$ is Theorem 2.1.3. $\supseteq$ is one line: $n = 2k$ gives $n^2 = 2(2k^2)$.

8.2.6 Induction, lemmas and symmetry

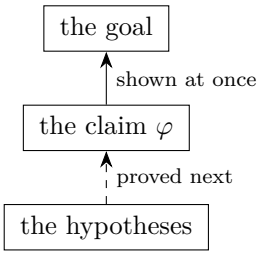
The rows of this table are not rules of the calculus. Each is a pattern of rules that proofs announce with a phrase of its own.

Pattern	Phrases	Silent in plain prose?
it suffices to show; a lemma (a cut, Lemma 2.3.3)	"we first show (Claim)"; "it suffices to show"; "we claim"	never the claim itself; "it suffices" may be dropped if the claim is displayed and then used
it suffices to prove something stronger	"it suffices to prove the stronger ..."; "in fact more is true: ..."	never the stronger claim; the step from it to the goal may be silent when it is one $\wedge$E, $\vee$I or chain link
a stronger hypothesis gives a weaker theorem	"in particular, for $x \in (0,1)$"; "since every such x has $\lvert x\rvert < 1$, the result applies"	the narrowing of the hypothesis, never
induction (Axiom 4.2.1, Theorem 4.2.2)	"we argue by induction on n"; "base case"; "suppose the claim holds for n"	never either the base case or the step; the induction hypothesis is stated

symmetry	"without loss of generality"; "similarly"; "by symmetry"	allowed only when the swap turns the statement and the hypotheses into themselves, up to identities such as $\lvert b-a\rvert = \lvert a-b\rvert$

8.2.6.1 It suffices to show

"It suffices to show φ" does two things. It shows on the spot that φ gives the goal, and it leaves φ to be proved: the goal is replaced by φ. As Section 1.3 observed, to show ψ it suffices to show φ when $\varphi \to \psi$ is in hand. A lemma is the same move in the other order: prove φ first, then use it as if it were a premise. (Logicians call both a **cut**; Lemma 2.3.3 (b) is the precise statement.) In analysis, "it suffices to show $|a_n - L| < \varepsilon$ for $n \geq N$" reduces a limit to one estimate; Example (The reciprocals tend to zero) made one with the claim $(*)$.



"It suffices to show φ": the solid step is written on the spot, the dashed one is the rest of the proof.

Quick check. In Example (The reciprocals tend to zero), which sentence is the "it suffices" step, and what does it leave to prove?

Solution. "It suffices to find $N \in \mathbb{Z}_{>0}$ with $1/N < \varepsilon$, since then …". It leaves $(*)$, which the next paragraph proves.

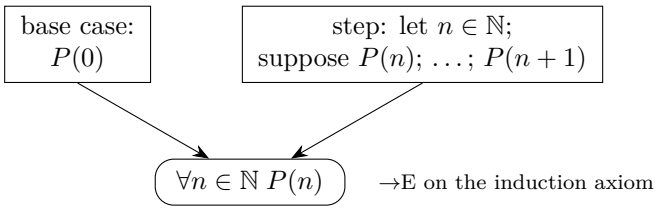
Two cases of the move have names of their own. To prove ψ it suffices to prove a **stronger** φ from which ψ follows at once: to show $x^2 < 1$ for $0 < x < 1$, show $x^2 < x$, and then $x^2 < x < 1$. And a theorem stays true when its hypothesis is narrowed: from "if $|x| < 1$, then $x^2 < 1$" and "every $x \in (0,1)$ has $|x| < 1$", the chain of Proposition 1.4.1 gives "if $x \in (0,1)$, then $x^2 < 1$": a stronger hypothesis, a weaker theorem.

Second reading · *Stronger and weaker, in general*

Both moves are instances of one fact about where a part sits in a formula. Weakening a part in a positive place, such as the conclusion of an implication, weakens the whole; weakening a part in a negative place, such as the hypothesis, strengthens the whole. The second-pass Section 4.9 states and proves this as Theorem 4.9.5; nothing on the main line uses it.

8.2.6.2 Induction

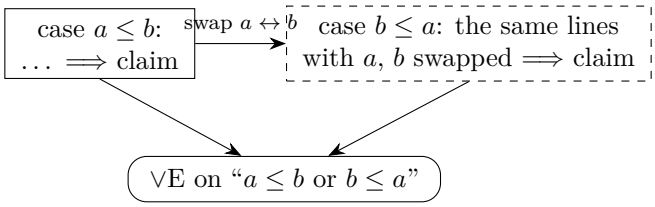
An induction is →E applied to Axiom 4.2.1, whose hypothesis is a conjunction: the base case $P(0)$, and the step "**for every** $n \in \mathbb{N}$, if $P(n)$ then $P(n+1)$". So the step is a ∀I box with a →I box inside it, and the phrase "suppose the claim holds for n" opens the inner box; its assumption is the **induction hypothesis**, and it must be stated, because the step uses it.



An induction: two inputs to one implication, the induction axiom (Axiom 4.2.1). The step is two boxes, the box of n and, inside it, the box of the induction hypothesis $P(n)$.

8.2.6.3 Without loss of generality

Symmetry is a case split ∨E whose second box is the first with two letters exchanged. "Without loss of generality $x \le y$" writes the first box and leaves the second to the reader, who must be able to get it by swapping the letters; that works when the statement and the hypotheses come back unchanged under the swap, up to identities such as $|b-a| = |a-b|$.



"Without loss of generality $a \le b$": the solid case is written, the dashed one is left to the reader, who gets it by the swap.

Quick check. May a proof that $x^2+y^2 \ge 2xy$ for all real x, y begin "Without loss of generality $x \le y$"?

Solution. Yes: swapping x and y leaves the statement unchanged, and there are no hypotheses. The case split is legitimate, though not needed, since $x^2 + y^2 - 2xy = (x-y)^2 \ge 0$.

Example. An induction, and a legitimate symmetry

(a) Let a be a real number with $0 < a \le 1$. Prove that $0 < a^n \le 1$ for every $n \in \mathbb{N}$, and annotate each phrase with its rule.

(b) Let a and b be real numbers. Prove that $|a-b| = \max\{a,b\} - \min\{a,b\}$, with a "without loss of generality", and write out the case split it abbreviates.

Solution. (a) For $n \in \mathbb{N}$, let $P(n)$ be the statement "$0 < a^n$ and $a^n \le 1$". We argue by induction on n, using Axiom 4.2.1.

Base case. $a^0 = 1$, and $0 < 1$ and $1 \le 1$, so $P(0)$ holds [∧I].

Step. Let $n \in \mathbb{N}$ [$\forall$I, opening], and suppose $P(n)$ holds, that is, $0 < a^n$ and $a^n \leq 1$ [$\to$I, opening: the induction hypothesis]. Then $a^{n+1} = a^n a$. Since $a^n > 0$ [$\wedge$E] and $a > 0$, the product $a^n a$ is positive, so $0 < a^{n+1}$. Since $a \leq 1$ and $a^n > 0$, multiplying gives $a^n a \leq a^n$, and $a^n \leq 1$ [$\wedge$E], so $a^{n+1} \leq 1$. Hence $P(n+1)$ holds [$\wedge$I].

By induction, $P(n)$ holds for every $n \in \mathbb{N}$ Axiom 4.2.1.

The base case is the first hypothesis of Axiom 4.2.1, written as its own paragraph. The step is the second: "let $n \in \mathbb{N}$" and "suppose $P(n)$ holds" open its two boxes, and "hence $P(n+1)$" is the last line of the inner one; the closing lines "so $P(n) \to P(n+1)$, for every n" are silent. The induction hypothesis is stated in full, and each of its halves is taken out by $\wedge$E and used: $a^n > 0$ twice (for the sign of a^{n+1}, and to multiply $a \leq 1$ by a^n without reversing it), and $a^n \leq 1$ once, at the end. The statement was chosen with care: the tempting "$a^n \leq a$" fails at $n = 0$ whenever $a < 1$, since $a^0 = 1$, and an induction for it would have to start at $n = 1$.

(b) Without loss of generality, $a \leq b$. Then $\max\{a, b\} = b$ and $\min\{a, b\} = a$, and since $b - a \geq 0$,

$$|a - b| = |b - a| = b - a = \max\{a, b\} - \min\{a, b\}.$$

The phrase abbreviates a $\vee$E on "$a \leq b$ or $b \leq a$", which holds for all real a, b by trichotomy. The first box is the paragraph above. The second box, on $b \leq a$, reads: then $\max\{a, b\} = a$ and $\min\{a, b\} = b$, and since $a - b \geq 0$, $|a - b| = a - b = \max\{a, b\} - \min\{a, b\}$. That is the first box with a and b exchanged throughout, read through three identities: $|b - a| = |a - b|$, $\max\{b, a\} = \max\{a, b\}$, and $\min\{b, a\} = \min\{a, b\}$. These say that the **statement** is unchanged when a and b are exchanged, and there are no hypotheses to check. This is what makes the swap exact, and so what licenses "without loss of generality".

Warning. "Without loss of generality" needs a statement and hypotheses that are unchanged by the swap. Here is a false "proof" that $\max\{a, b\} = b$ for all real a, b: "Without loss of generality $a \leq b$; then $\max\{a, b\} = b$." Exchanging a and b turns the claim into $\max\{b, a\} = a$, which is a different claim, so the second case is not the first one in disguise; and indeed $a = 2$, $b = 1$ gives $\max\{a, b\} = 2 \neq 1$. Before writing "without loss of generality", swap the letters in the statement and the hypotheses, and check that each comes back unchanged, up to identities such as $|b - a| = |a - b|$.

8.2.7 What the sentence does not say

Some quantifiers in mathematical English are not written at all, and some are written where they do not apply. Four habits of mathematical English cover most of them.

The hidden universal closure. A theorem stated as "if n is even, then n^2 is even" means "for every integer n, if n is even, then n^2 is even". Section 5.1 introduced this convention: *a letter that is not introduced anywhere, in the statement of a theorem, is read with a "for every" in front of the whole statement.* It applies to the statement of a theorem, lemma or exercise, not to a line inside a proof: inside a proof every letter has been introduced by a "let" or a "choose", and its box tells the reader which. The proof of such a theorem begins by opening the hidden box: "let n be an integer".

"For all large n" hides an $\exists N$ and a bound, as the Quick check above found. "Eventually" means the same. The negation of "$P(n)$ for all large n" is "for every N there is an $n \geq N$ for

which $P(n)$ fails", usually said "$P(n)$ fails for infinitely many n".

"Any" can mean "every" or "some", and only the rest of the sentence decides (Section 5.1). "Any $x \in A$ is positive" means every element of A is positive. "If any $x \in A$ is positive, then A is non-empty" means: if **some** element of A is positive. When "any" appears, rewrite the sentence once with "every" and once with "some", and keep the one that was meant.

Quantifiers at the end of a sentence. English lets a writer put "for some x" or "for all y" after the formula they govern, and then the order of two such phrases is unclear. "$x^2 > y$ for some x for all y", over $\mathbb{R}$, has two readings:

$$\forall y\, \exists x\ (x^2 > y) \qquad \text{and} \qquad \exists x\, \forall y\ (x^2 > y).$$

The first is true (for $y \geq 0$ take $x = y + 1$, since $(y+1)^2 = y^2 + 2y + 1 > y$; for $y < 0$ take $x = 0$). The second is false (whatever x is proposed, $y = x^2$ refutes it). The remedy is to write the quantifiers of such a statement **in front**, in the order in which they apply. A trailing phrase such as "for every $n \geq N$" is safe only when it is the last quantifier of the sentence, as in the definition of $a_n \to L$.

Warning. Moving "for all x" to the end of a sentence can change continuity into uniform continuity. "$|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$, for all x", written after "for every $\varepsilon > 0$ there is $\delta > 0$ such that", does not say whether $\forall x$ comes before $\exists \delta$ (continuity at every point) or after it (uniform continuity). The two formulas of Section 6.4 differ exactly there, and Example (The square function on the real line) shows that the difference is real: $f(x) = x^2$ on $\mathbb{R}$ satisfies the first and not the second. Write the quantifiers in front.

Warning (Not allowed: quantifier steps in prose). Each sentence below might appear in a proof, and each is wrong. The tag names the entry of the catalogue in Section 8.6, or the item of the convention on silent steps in Section 8.3, that it breaks.

- "Let $\varepsilon > 0$ and choose N with $|a_n - L| < \varepsilon$ for every $n \geq N$. This N works for every $\varepsilon > 0$." The N was chosen inside the box of one ε: (N2).
- "Some $x \in A$ has $f(x) > 0$. Let $x \in A$; then $f(x) > 0$." The witness of an existential is *chosen* ("choose $x \in A$ with $f(x) > 0$", $\exists$E with a fresh name). Written with "let", the second sentence claims $f(x) > 0$ for every $x \in A$: (N3).
- "Choose N with $|a_n - L| < \varepsilon/2$ for $n \geq N$, and N with $|b_n - M| < \varepsilon/2$ for $n \geq N$." One name for two witnesses: (N4).
- "Let L be the limit of (a_n)", in a proof that has not shown that (a_n) converges: (N14).
- "Let $\varepsilon > 0$. Then $|a_n - L| < \varepsilon$ for every $n \geq N$, as required", where no sentence says what N is. The witness of $\exists$I, with what it depends on, is never silent (Definition 8.3.2, Section 8.3).

8.2.8 Exercises

8.2.8.1 A. Check your understanding

Exercise. A1

Let (a_n) be a sequence of real numbers and L a real number. Here is a proof that if $a_n \to L$ and $L > 0$, then $a_n > 0$ for all large n. Its sentences are numbered.

(1) Suppose $a_n \to L$ and $L > 0$. (2) Then $L/2 > 0$. (3) Choose $N \in \mathbb{N}$ such that $|a_n - L| < L/2$ for every $n \geq N$. (4) Let $n \in \mathbb{N}$ with $n \geq N$. (5) Then $L - a_n \leq |a_n - L| < L/2$, so $a_n > L/2 > 0$. (6) Hence $a_n > 0$ for every $n \geq N$, that is, $a_n > 0$ for all large n.

(a) Name the rule or rules carried by each sentence.

(b) Sentence (3) has no dependency parenthesis. What should it say, and why does N not depend on n?

(c) Which instance of ∀E in the proof is not the obvious one?

(d) Which closing lines are silent?

Solution. (a) (1) →I, opening: the hypothesis of the theorem is assumed (its two conjuncts are then available by ∧E). (2) An arithmetic step from $L > 0$. (3) ∀E on the definition of $a_n \to L$ at $L/2$, with its bound from (2), then ∃E, opening a box with the fresh name N. (4) ∀I and →I, opening the boxes of n and of $n \geq N$. (5) ∀E at n and →E on the line from (3), giving $|a_n - L| < L/2$; then the school fact $-u \leq |u|$ with $u = a_n - L$, and arithmetic. (6) →I and ∀I, closing ("for every $n \geq N$"), then ∃I with the witness N ("for all large n" is $\exists N\, \forall n \geq N$), and the ∃E box of (3) closes, since the conclusion does not mention N.

(b) "Choose $N \in \mathbb{N}$ (depending on L) such that …". N is obtained at line (3), from the definition instantiated at $L/2$, a term that mentions L (and the sequence, which is data). The box of n is opened only at (4), after N is chosen, so N cannot depend on n.

(c) The instance $\varepsilon = L/2$ in (3). It is the instance the proof is built around: any ε with $0 < \varepsilon \leq L$ would do, and $L/2$ leaves the margin $a_n > L/2$.

(d) The closing lines of the ∃E box (3) and of the →I box (1), "hence, if $a_n \to L$ and $L > 0$, then …". Sentence (6) writes the closings of the boxes of (4) in words, and the ∃I through "that is".

Exercise. A2: Drill

Each part takes a line or two. Most parts ask which rule a phrase carries, whether it may be silent, or for a judgment or a sentence.

(a) "Fix $\varepsilon > 0$." Which rule, and may it be silent?

(b) "Applying the hypothesis with $\varepsilon/3$ in place of ε." Which rule, and may the term go unnamed?

(c) "Take $\delta := \varepsilon/5$." Which rule, may it be silent, and what does δ depend on?

(d) "By hypothesis there is k with $m = 2k + 1$; choose such a k." Which rule, and what does it require of the name?

(e) In "Let $\varepsilon > 0$. Choose N such that … . Let $n \geq N$.", may N depend on n?

(f) In "Let $x \in A$. Let $\varepsilon > 0$. Choose $\delta > 0$ such that …", may δ depend on x?

(g) Write "$x^2 > 0$ for all $x \neq 0$" with the quantifier in front.

(h) Negate "$a_n > 0$ for all large n" with no ¬ in front of a quantifier.

(i) A proof of "$\max\{a, b\} \geq a$" begins "Without loss of generality $a \leq b$." Is that legitimate?

(j) In an induction, which sentence opens the box of the induction hypothesis?

(k) To prove $A = B$, a proof shows $A \subseteq B$ and $B \subseteq A$. Which rules does that pattern use, and how does each inclusion open?

(l) A proof begins "Let m be the least element of A". What must already be on the page?

Solution.

(a) Bounded $\forall$I, opening; never silent.

(b) $\forall$E at a term that is not the obvious one; the term must be named.

(c) $\exists$I's witness, written before the work; never silent; it depends on ε.

(d) $\exists$E, with a fresh name.

(e) No: the box of n opens after N is chosen.

(f) It may, since both boxes are open; whether it does depends on the line it comes from.

(g) $\forall x\,(x \neq 0 \to x^2 > 0)$.

(h) For every N there is $n \geq N$ with $a_n \leq 0$.

(i) No: the swap turns the claim into $\max\{b, a\} \geq b$, a different claim; each case needs its own line (both are one line long).

(j) "Suppose $P(n)$ holds", inside "Let $n \in \mathbb{N}$".

(k) $\leftrightarrow$I inside $\forall$I, for $\forall x\,(x \in A \leftrightarrow x \in B)$; then $A = B$ by $\forall$E and $\to$E with the axiom of extensionality, cited. Each inclusion is a bounded $\forall$I that opens with "let $x \in A$" (or "let $x \in B$").

(l) That such an element exists and that there is only one ($\exists!$, a definite description). Uniqueness is one line, since two least elements are each at most the other; existence is the part that needs a reason.

8.2.8.2 B. Practice

Exercise. B1: Ambiguous English

Each sentence has a quantifier phrase at its end. Write both readings in symbols, with the quantifiers in front, and decide whether each reading is true for the function given.

(a) "$f(x) \leq M$ for every real x, for some real M", for $f(x) = x$.

(b) "$x + y = 0$ for some real y, for every real x."

(c) "$|x - y| < \delta$ implies $|f(x) - f(y)| < 1$ for all real x, y, for some $\delta > 0$", for $f(x) = x^2$.

Solution. *(a)* $\exists M\,\forall x\ (f(x) \leq M)$, "$f$ is bounded above", and $\forall x\,\exists M\ (f(x) \leq M)$. For $f(x) = x$ the first is false: whatever M is proposed, $x = M + 1$ has $f(x) = M + 1 > M$. The second is true for every function: given x, take $M = f(x)$.

(b) $\forall x\,\exists y\ (x + y = 0)$ and $\exists y\,\forall x\ (x + y = 0)$. The first is true: given x, take $y = -x$. The second is false: for any y, the two values $x = -y$ and $x = 1 - y$ give $x + y = 0$ and $x + y = 1$, so not every x works.

(c) $\exists \delta > 0\ \forall x \forall y\ (|x - y| < \delta \to |f(x) - f(y)| < 1)$ and $\forall x \forall y\ \exists \delta > 0\ (|x - y| < \delta \to |f(x) - f(y)| < 1)$. For $f(x) = x^2$ the first is false: it is the uniform continuity condition

at $\varepsilon = 1$, and Example (The square function on the real line) refutes it with $x = 1/\delta$, $y = 1/\delta + \delta/2$, for which $|x - y| < \delta$ and $|x^2 - y^2| = 1 + \delta^2/4 \geq 1$. The second is true for every function: given x and y, if $x = y$ then $|f(x) - f(y)| = 0 < 1$ for any δ, say $\delta = 1$; if $x \neq y$, take $\delta = |x - y|$, so that $|x - y| < \delta$ is false and the implication holds vacuously.

In (a) and (c), the reading with $\exists$ in front is the one a writer means; the other is true of every function, so it says nothing about f. In (b), the intended reading is $\forall x\, \exists y$. Written with the quantifiers in front, the sentences say what they mean: "there is M such that $f(x) \leq M$ for every x", "for every x there is y with $x + y = 0$", "there is $\delta > 0$ such that for all $x, y, \ldots$".

Exercise. B2: An induction, annotated

Prove that $5^n - 1$ is divisible by 4 for every $n \in \mathbb{N}$. Annotate each phrase of your proof with its rule, and say which hypothesis of Axiom 4.2.1 each part of the proof establishes.

Solution. For $n \in \mathbb{N}$, let $P(n)$ be "$4 \mid 5^n - 1$", that is, "there is an integer k with $5^n - 1 = 4k$". We argue by induction on n.

Base case. $5^0 - 1 = 0 = 4 \cdot 0$, so $P(0)$ holds, with the witness $k = 0$ [$\exists$I]. This is hypothesis (i) of Axiom 4.2.1.

Step. Let $n \in \mathbb{N}$ [$\forall$I, opening], and suppose $P(n)$ holds [$\to$I, opening: the induction hypothesis]. By the induction hypothesis there is an integer k with $5^n - 1 = 4k$; choose such a k (depending on n) [$\exists$E, fresh name k]. Then

$$5^{n+1} - 1 = 5 \cdot 5^n - 1 = 5(5^n - 1) + 4 = 5 \cdot 4k + 4 = 4(5k + 1),$$

where the third equality substitutes $5^n - 1 = 4k$ [=E]. Since $5k + 1$ is an integer, $P(n + 1)$ holds, with the witness $5k + 1$ [$\exists$I]. The step is hypothesis (ii) of Axiom 4.2.1: the $\forall$I on n and the $\to$I on $P(n)$.

By induction, $4 \mid 5^n - 1$ for every $n \in \mathbb{N}$ Axiom 4.2.1.

The $\exists$E box of k closes on "$P(n+1)$", which does not mention k, as its side condition requires; its closing line, and those of the two boxes of the step, are silent.

Exercise. B3: Six, two and three, used later

Let $A \coloneqq \{n \in \mathbb{Z} : 6 \mid n\}$ and $B \coloneqq \{n \in \mathbb{Z} : 2 \mid n \text{ and } 3 \mid n\}$. Prove $A = B$ by double inclusion. Name each "let", each "choose" with its fresh name, and each witness.

Solution.

- ($\subseteq$) Let $n \in A$ [$\forall$I], and choose k with $n = 6k$ [$\exists$E, k fresh]. Then $n = 2(3k)$ and $n = 3(2k)$, so $2 \mid n$ and $3 \mid n$, with the witnesses $3k$ and $2k$ [$\exists$I twice, $\wedge$I]; so $n \in B$.
- ($\supseteq$) Let $n \in B$ [$\forall$I]. Choose a with $n = 2a$ and b with $n = 3b$ [$\exists$E twice, two fresh names: one name for both would be (N4)]. Then $b = 3b - 2b = n - 2b = 2a - 2b = 2(a - b)$, so $n = 3b = 6(a - b)$, and $6 \mid n$ with the witness $a - b$ [$\exists$I]; so $n \in A$.
- So every integer n has $n \in A \leftrightarrow n \in B$ [$\leftrightarrow$I inside $\forall$I], and by extensionality $A = B$.

The two ∃E boxes of (⊇) close on "$n \in A$", which mentions neither a nor b.

Marking scheme. (10 points) - 2: (⊆), with k chosen and both witnesses written. - 4: (⊇): two choices with distinct names (1); the computation $b = 2(a - b)$, or another valid route to an integer c with $n = 6c$ (2); the witness named (1). - 2: the double-inclusion frame: both inclusions, each opened by "let $n \in \ldots$", and the conclusion $A = B$ by extensionality, cited. - 2: the tags: ∀I, ∃E with fresh names, ∃I.

Fatal errors (0 for the step): one name for both witnesses; "$2 \mid n$ and $3 \mid n$, so $6 \mid n$" asserted without an argument (N9).

8.2.8.3 C. Going deeper

Exercise. C1: When "similarly" is honest

(a) Prove that $\max\{x, y\} + \min\{x, y\} = x + y$ for all real x and y, with a "without loss of generality". Then write out the ∨E it abbreviates, and check that its second box is the first with the letters exchanged.

(b) Let $x \geq 0$ and y be real numbers. A student proves $x \max\{x, y\} \geq xy$ as follows: "Without loss of generality $x \leq y$. Then $\max\{x, y\} = y$, and $xy \geq xy$." Explain why "without loss of generality" is not legitimate here, prove the case the student skipped, and show by an example that the hypothesis $x \geq 0$ cannot be dropped from the statement.

Solution (Hint 1). In (b), exchange x and y in the statement and in the hypothesis, and compare with the original.

Solution (Hint 2). In the case $y \leq x$, $\max\{x, y\} = x$, and the claim becomes $x^2 \geq xy$.

Solution. (a) Without loss of generality $x \leq y$. Then $\max\{x, y\} = y$ and $\min\{x, y\} = x$, so $\max\{x, y\} + \min\{x, y\} = y + x = x + y$.

The phrase abbreviates ∨E on "$x \leq y$ or $y \leq x$", true for all real x, y by trichotomy. Box 1, on $x \leq y$, is the paragraph above. Box 2, on $y \leq x$: then $\max\{x, y\} = x$ and $\min\{x, y\} = y$, so $\max\{x, y\} + \min\{x, y\} = x + y$. Exchanging x and y in box 1 gives "$y \leq x$; $\max\{y, x\} = x$, $\min\{y, x\} = y$; $\max\{y, x\} + \min\{y, x\} = x + y = y + x$", which is box 2 read through $\max\{y, x\} = \max\{x, y\}$, $\min\{y, x\} = \min\{x, y\}$ and $y + x = x + y$. These identities say that the statement is unchanged by the exchange, and there are no hypotheses, so the swap is exact.

(b) Exchanging x and y changes both the hypothesis and the claim: it gives "$y \geq 0$, and $y \max\{y, x\} \geq yx$", which is not the original statement. So the case $y \leq x$ is not the case $x \leq y$ in disguise, and it needs its own argument.

Case $y \leq x$. Then $\max\{x, y\} = x$, and multiplying $x \geq y$ by $x \geq 0$ gives $x \max\{x, y\} = x^2 \geq xy$.

With the student's case ($x \leq y$: $x \max\{x, y\} = xy \geq xy$), this proves the claim in either

case. The hypothesis $x \geq 0$ is used only in the skipped case. It cannot be dropped: $x = -1$, $y = -2$ has $\max\{x, y\} = -1$, so $x \max\{x, y\} = 1$, while $xy = 2$, and $1 \geq 2$ is false.

Marking scheme. (10 points) - 2: (a) a correct computation in the case $x \leq y$. - 3: (a) the $\vee$E written out, with its disjunction justified, and the exchange checked against the statement. - 2: (b) the reason: the exchange changes the hypothesis and the claim. - 2: (b) the case $y \leq x$ proved, using $x \geq 0$. - 1: (b) a counterexample without $x \geq 0$, checked. Fatal errors (0 for the step): in (b), "the other case is similar" without computing it; a counterexample that satisfies $x \geq 0$.

Exercise. C2: Moving a quantifier changes the meaning

Give an example of a sentence of analysis, in English, that begins "for every x ...", whose meaning changes when the phrase "for every x" is moved to the end of the sentence. Write both readings in symbols and show that one is true and the other false.

Solution. An example. "For every real $x > 0$, there is $n \in \mathbb{Z}_{>0}$ with $1/n < x$." Moved: "There is $n \in \mathbb{Z}_{>0}$ with $1/n < x$, for every real $x > 0$." The readings are

$$\forall x > 0\ \exists n \in \mathbb{Z}_{>0}\ (1/n < x) \qquad \text{and} \qquad \exists n \in \mathbb{Z}_{>0}\ \forall x > 0\ (1/n < x).$$

The first is true: given $x > 0$, there is an integer $n > 1/x$ (the school fact of Example (The reciprocals tend to zero)), and $n > 1/x > 0$ gives $n \in \mathbb{Z}_{>0}$ and $1/n < x$. The second is false: whatever $n \in \mathbb{Z}_{>0}$ is proposed, $x = 1/n$ is positive and $1/n < x$ fails. (It is natural to read the moved sentence as the second formula, since its "for every x" now comes after "there is n"; whichever reading a reader picks, the sentence no longer says one thing.)

What a different answer must have, so that it can be checked against this one: - the sentence contains "for every x" and at least one existential quantifier ("there is", "for some") whose scope contains a formula mentioning x; - moved to the end, "for every x" can be read as coming after the existential quantifier, so that the witness may no longer depend on x; - both readings are written in symbols with all quantifiers in front, and there is a specific interpretation (here $\mathbb{R}$ with its usual order) in which one reading is true and the other false, each claim checked; - the false reading is refuted by naming, for each candidate witness, a value of x that defeats it. Another correct answer, for $f(x) = x^2$ on $\mathbb{R}$, is "for every x, for every $\varepsilon > 0$ there is $\delta > 0$ such that $|f(x) - f(y)| < \varepsilon$ whenever $|x - y| < \delta$" (continuity, true); with "for every x" moved to the end, it becomes the sentence of the warning, whose uniform-continuity reading Example (The square function on the real line) refutes.

8.3 Four Levels of One Proof

Sections 8.1 and 8.2 matched each rule with the phrases that carry it. A proof, though, is read and written whole. This section writes proofs several times, from a full derivation to the plain prose of a textbook, names the four levels this gives, fixes exactly which steps the last level may leave unsaid, and ends with the use that matters most to a reader working alone: when a sentence looks doubtful, rewrite it, and only it, as a skeleton, and look.

8.3.1 One short proof, four texts

Here is one proof, for every integer n: *if n is odd, then n^2 is odd*, written four times, starting from the version a textbook prints and zooming in.

Plain prose.

> Let n be odd, say $n = 2k + 1$. Then $n^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$, so n^2 is odd.

Prose with the rules tagged.

> Let n be an integer [$\forall$I], and suppose n is odd [$\to$I]; choose an integer k with $n = 2k+1$ [$\exists$E, fresh k]. Then $n^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$ [algebra], so n^2 is odd, with the witness $2k^2 + 2k$ [$\exists$I; then $\exists$E closes, since "n^2 is odd" does not mention k]. Since n was arbitrary, every odd integer has an odd square [$\to$I, $\forall$I, closing].

The skeleton (domain $\mathbb{Z}$).

1	n_0	$\exists k\,(n_0 = 2 \cdot k + 1)$	asm, for $\to$I
2		k_0 | $n_0 = 2 \cdot k_0 + 1$	asm, for $\exists$E
3		$n_0 \cdot n_0 = 2 \cdot (2 \cdot k_0 \cdot k_0 + 2 \cdot k_0) + 1$	arith 2
4		$\exists j\,(n_0 \cdot n_0 = 2 \cdot j + 1)$	$\exists$I 3
5		$\exists j\,(n_0 \cdot n_0 = 2 \cdot j + 1)$	$\exists$E 1, 2–4
6		$\exists k\,(n_0 = 2 \cdot k + 1) \to \exists j\,(n_0 \cdot n_0 = 2 \cdot j + 1)$	$\to$I 1–5
7		$\forall n\,(\exists k\,(n = 2 \cdot k + 1) \to \exists j\,(n \cdot n = 2 \cdot j + 1))$	$\forall$I 1–6

The full derivation. The skeleton again, with its line 3 split into two arith lines, each recording one fact: the square of $2k_0 + 1$, then a factor 2 taken out.

1	n_0	$\exists k\,(n_0 = 2 \cdot k + 1)$	asm, for $\to$I
2		k_0 | $n_0 = 2 \cdot k_0 + 1$	asm, for $\exists$E
3		$n_0 \cdot n_0 = 4 \cdot k_0 \cdot k_0 + 4 \cdot k_0 + 1$	arith 2
4		$n_0 \cdot n_0 = 2 \cdot (2 \cdot k_0 \cdot k_0 + 2 \cdot k_0) + 1$	arith 3
5		$\exists j\,(n_0 \cdot n_0 = 2 \cdot j + 1)$	$\exists$I 4
6		$\exists j\,(n_0 \cdot n_0 = 2 \cdot j + 1)$	$\exists$E 1, 2–5
7		$\exists k\,(n_0 = 2 \cdot k + 1) \to \exists j\,(n_0 \cdot n_0 = 2 \cdot j + 1)$	$\to$I 1–6
8		$\forall n\,(\exists k\,(n = 2 \cdot k + 1) \to \exists j\,(n \cdot n = 2 \cdot j + 1))$	$\forall$I 1–7

Four texts, one proof. Each keeps the box of n, the assumption that n is odd, the choice of k and the computation; they differ in how much of the routine work the reader is trusted to restore.

A derivation, a skeleton and a paragraph are the same proof; they differ in what the reader is trusted to restore.

Quick check. In the plain prose, which words stand for line 2 of the skeleton, and which skeleton lines have no words at all?

Solution. "Say $n = 2k + 1$" is line 2, the opening of the ∃E box with the fresh name k. Lines 5, 6 and 7, the closings of the three boxes, have no words. "So n^2 is odd" is line 4, the ∃I.

Sections 2.2 and 6.4 already printed proofs twice, once as a derivation and once as a paragraph. The four texts you have just read are the four levels of the next definition.

Definition 8.3.1. Levels of Proof

A proof can be written at four **levels**.

- **Level 0, the full derivation.** A derivation in the sense of Definition 1.4.2, with the rules of Chapters 1, 2 and 5: **every line** is a premise, an assumption, or **one** rule applied to available lines. A derived rule used as one line (Definition 2.3.1), such as the bounded rules of Proposition 6.1.1, counts as one rule. Facts about numbers enter as premises, or as lines justified "arith", each of which records **one** such fact applied to the lines it cites.
- **Level 1, the skeleton.** The boxes and the logical lines only: every assumption, every flagged name, every line that opens or closes a box, with the rule that closes it, and every line that such a closing line cites. Each stretch of mathematics between them is compressed into one line, justified "arith", "by def", or by a cited result; a stretch of routine logical steps inside one box is compressed into one line justified by the names of its rules in order, such as "∀E, ∃I".
- **Level 2, annotated prose.** Sentences, each tagged in brackets with the rule or rules it carries, such as [∀I] or [∃E].
- **Level 3, plain prose.** Sentences without tags, in which some steps are left unsaid. Which steps may be left unsaid is fixed by the convention that follows, not by the writer.

In words: level 0 hides no logic. Every line other than an arith line can be checked mechanically against the schema of its rule, and each arith line records one fact about numbers that can be checked by hand. Level 1 keeps the logical architecture and lets the arithmetic run in blocks. Level 2 is level 1 turned into sentences, with the architecture still labeled. Level 3 removes the labels and some of the lines, following a fixed convention.

Second reading · *Levels already on the page, and a skeleton missing a box*

The levels are already on the page in earlier chapters. The derivation in Section 6.4 is level 0. The paragraph after Theorem 2.2.1, with its tags [→E], [∧I], [¬E], [¬I], is level 2 for the rules Chapter 2 had; its quantifier steps carry no tags. The derivation printed below that paragraph is not quite level 0, and the reason is instructive: it fixes r, p and q in advance and leaves out the two quantifier steps, choosing p, q (and later m) and passing from one r to every r, because Chapter 2 had no rules for them yet. Section 2.2 says so in the paragraph after the derivation. With Chapter 5's rules those steps can be written. The example of the square root of two below writes the choices of p and q and the step to every r; the choice of m stays inside the arithmetic

fact of line 15 of its level-0 derivation (in the fold "Level 0 of the square root of two"), since evenness is treated there as a single predicate E.

A **non-example**, by one change: take a level-1 skeleton and delete one of its boxes, say the box in which a name N_0 is chosen, keeping the lines inside. The result is not a shorter skeleton. It is a list in which N_0 appears from nowhere, and no reader can tell whether it was chosen, given, or assumed. A skeleton may compress the work inside a box; it may not compress the boxes.

Why four levels and not two? Because each serves a task. Level 0 is for settling a doubt completely: nothing in it is taken on trust. Level 1 is for planning a proof and for reading one: it is what you write on scratch paper before the first sentence (Section 8.5 turns this into a procedure), and what you reconstruct from a textbook when a proof will not come clear (Section 8.4). Level 2 is for learning, which is why this volume uses it. Level 3 is what mathematicians read and write.

Warning. **A skeleton is not a finished proof.** A level-1 text with lines such as "$|x_0 - y| < \delta_0 \to |f(x_0) - f(y)| < \varepsilon_0$, arith" is a plan: it says that a computation goes there, and it can hide a computation that does not work. For $f(x) = x^2$ on $\mathbb{R}$, a skeleton of "f is uniformly continuous" with one line "arith" at that point looks exactly like a correct skeleton for $f(x) = 3x + 1$, where $\delta = \varepsilon/3$ works because $|(3x+1) - (3y+1)| = 3|x - y|$; the claim is false for the first function (Example (The square function on the real line)), and only writing out the arithmetic shows where it fails. Expand every "arith" line before calling the proof done.

8.3.2 Which steps may be silent

Section 8.1 stated the principle: a step may go unsaid when its result is forced by the lines around it; every choice, assumption and citation is spoken. The most common cases:

The step	Silent?	Why
the closing line of $\to$I, $\forall$I, $\vee$E or $\exists$E (not of $\neg$I, RAA or $\leftrightarrow$I): "since x was arbitrary", "in either case", "hence if P then Q"	may be	the box decides it
$\wedge$E; $\vee$I; repeating a line already shown	may be	the next sentence shows which part or form was used
$\forall$E at the name in sight; =E inside a chain	may be	the sentence that uses it forces it
opening a box: "let", "suppose", "choose ..." with a new name	never	it creates a name or an assumption
a witness, with what it depends on; a case split, with why the cases cover everything	never	it is a choice

The step	Silent?	Why
a citation of a hypothesis or a result; the clash of ¬E; the conclusion of ¬I or RAA	never	the reader cannot guess what licensed it

The official list, which the rest of the book checks written proofs against, is the next definition. The table gives its most common entries; the definition gives all of them.

Definition 8.3.2. Silent Steps

In level-3 prose the following steps **may be silent**, that is, left unwritten:

- ∧I when both conjuncts are displayed, and ∧E;
- R;
- ∨I;
- ⊥E, once the contradiction or the empty case has been stated;
- the **name** of DS (not the fact it uses);
- the exchange of the two members of a disjunction or of a conjunction (Example (First derivations with "or"), Proposition 1.2.1), made only to match the schema of a rule, with no citation;
- ↔E with a definition ("that is", "by definition", or nothing);
- ∀E at the obvious term;
- →E whose implication is a line already in the proof that is neither a hypothesis of the statement, nor a cited result, nor an instance of one (for instance, an instance of the property of a chosen name), when its antecedent is displayed or has just been assumed;
- =I and =E inside a chain of equalities or inequalities;
- the phrase "it suffices to show", when the claim of the cut is displayed and then used;
- the **closing** lines of →I ("hence if P then Q"), of ∀I ("since x was arbitrary"), of ∨E ("in either case") and of ∃E (the conclusion repeated outside the box of the chosen name).

The following are **never silent**:

- the openings of →I, ¬I, RAA and ∀I, and of ∃E **with a fresh name**;
- the conclusions of ¬I and of RAA ("hence ¬P", "hence P");
- a ¬¬E or LEM step, written as an explicit case split or as an explicit removal of a double negation (Proposition 2.5.1 (b); "P must hold, since ..."), with the flag of Section 8.7 when the step is genuinely non-constructive;
- the case split of ∨E, with the reason the cases exhaust;
- the clash of ¬E: both members of the clashing pair are named;
- the witness of ∃I, with what it depends on;
- an instance of ∀E at a term that is not the obvious one;
- every citation, and so every →E whose implication is a hypothesis of the statement or a cited result, or an instance of one;
- both directions of ↔I, and both halves of an ∃! proof;
- the base case and the step of an induction, with the induction hypothesis stated;
- the announcement of a contrapositive;
- the negated form, after De Morgan (Proposition 2.5.5) or a quantifier negation (Proposition 6.2.1, Proposition 6.2.3);

- the claim of a cut;
- a "without loss of generality", together with the symmetry that licenses it.

The eighth and ninth items of the first list, for instance, say this: after "choose N such that $|a_n - L| < \varepsilon$ for every $n \geq N$", applying that line to a particular $n \geq N$ needs no comment.

Silent steps and truth tables. Section 4.8 named the step that a truth table licenses, TAUT (Definition 4.8.1). Section 6.3 let it act on the propositional skeleton of a first-order line, where every atomic and every quantified part counts as a letter (Definition 6.3.1, Proposition 6.3.1). Most items of the first list are TAUT steps: $\wedge$I with both conjuncts displayed, $\wedge$E, R, $\vee$I, $\bot$E once the contradiction is stated, DS, the exchange of the members of a disjunction or conjunction, $\leftrightarrow$E with a definition, and $\rightarrow$E on a line of the proof. A reader can check each of them by a truth table on the skeleton. The quantifier and equality items, $\forall$E at the obvious term and =I and =E inside chains, are not TAUT steps: they are silent because the sentence that uses them forces them. TAUT never licenses a first-order law. On the skeleton, $\forall x\, \forall y\, R(x, y)$ and $\forall y\, \forall x\, R(x, y)$ are two different letters, so passing from one to the other is a derivation with the quantifier rules, not a TAUT step. From Section 6.3 on, a level-1 line that compresses propositional steps inside one box may cite TAUT with its lines, in place of the names of its rules.

Warning. **Level 3 is not level 0 with steps missing at random.** A proof that skips the reason a case split is exhaustive, or chooses δ without saying what it depends on, or uses a lemma without naming it, is not "written at level 3": it has a gap, because those steps are in the never-silent list. Conversely, a proof that omits "since ε was arbitrary" has no gap. When you judge whether a written proof is complete, judge it against Definition 8.3.2, not against your sense of what is obvious.

↻ **Second reading** · *The two lists against the principle, with an example*

Read the two lists against the principle. Everything in the first list is forced. $\wedge$E has one premise and two possible conclusions, and the next sentence shows which was used. The closing line of a box is determined by the box. $\forall$E at the obvious term x, when x is the name in sight, is forced by the sentence that uses it. Everything in the second list carries information that the surrounding lines do not determine. Which name is chosen, and that it is fresh; which witness, and what it depends on; which cases, and why they are all the cases; which result is cited. A reader who meets "Then p is even" with no citation has to guess what licensed it, and a guess is not a check.

A small example. In "since $x \in (0, 1)$, in particular $x > 0$", the step from the conjunction "$0 < x$ and $x < 1$" to its first conjunct is $\wedge$E, and the unwinding of $x \in (0, 1)$ into that conjunction is $\leftrightarrow$E with the definition of an interval. Both may be silent, and here the writer kept the phrase "in particular" anyway, which is allowed: the convention says what **may** be omitted, never what must be. On the other side, "suppose $r^2 = 2$" can never be omitted from the proof of Theorem 2.2.1: without it, the reader does not know that a box has opened, and the final "$r^2 \neq 2$" has no visible source.

Quick check. At level 3, may each of these be left out: (i) "since ε was arbitrary"; (ii) "choose $\delta > 0$ (depending on ε) such that ..."; (iii) "by the triangle inequality"?

Solution.

(i) Yes: it is the closing of ∀I. (ii) No: it is an ∃E with a fresh name, and the dependence of a witness. (iii) No: it is a citation.

8.3.3 The irrationality of the square root of two, four times

The first example is the proof of Theorem 2.2.1, now with every quantifier step written. It uses the fact that Section 2.2 calls (L): every rational number is p/q for integers p, q with $q \neq 0$, not both even. Write $E(x)$ for "x is even". The proof uses only two clauses of (L), and in a convenient form: $r = p/q$ with $q \neq 0$ gives $p = r \cdot q$, which avoids division; and "not both even" is $\neg(E(p) \wedge E(q))$. The clause $q \neq 0$ is then not needed at all, since $q = 0$ would force $p = 0$, and both would be even. It also uses Theorem 2.1.3, which we write as the formula $\forall n \in \mathbb{Z}\ (E(n \cdot n) \to E(n))$. To keep the lines short, write $L(r, p, q)$ for $p = r \cdot q \wedge \neg(E(p) \wedge E(q))$.

Example. The irrationality of the square root of two at four levels

Write the proof of Theorem 2.2.1, that no rational number has square 2, at levels 0, 1, 2 and 3 of Definition 8.3.1, and match the levels line by line.

Idea. The move this example shows in full is **skeleton first**: before any arithmetic, write the first and the last line of every box. The goal $\forall r \in \mathbb{Q}\ \neg(r \cdot r = 2)$ dictates the outer box (∀I, "let $r \in \mathbb{Q}$") and the next one (¬I, "suppose $r^2 = 2$", ending in $\bot$). Inside, (L) instantiated at r_0 is an existential, so its ∃E boxes open at once, with fresh names p_0, q_0, and they too end in $\bot$. Only then is there something to compute: the arithmetic fills the innermost box, and it has a target, the clash with "not both even". Sections 8.5 and 8.6 come back to this move twice, with parts of the proof left blank.

Before reading on: the level-3 proof below has four sentences. Which lines of the level-1 skeleton does its second sentence, "By (L), write $r = p/q$ ...", stand for?

Solution. **Level 3.**

> Let r be rational, and suppose $r^2 = 2$. By (L), write $r = p/q$ with integers p, q, not both even. Then $p^2 = 2q^2$, so p^2 is even, and p is even by Theorem 2.1.3; writing $p = 2m$ gives $q^2 = 2m^2$, so q is even by the same theorem. Thus p and q are both even, a contradiction; hence $r^2 \neq 2$.

Level 2.

> Let $r \in \mathbb{Q}$ [∀I: the box of r_0], and suppose $r^2 = 2$ [assumption for ¬I]. By (L), there are integers p and q with $p = rq$, not both even; choose such p and q [∀E at

r, then ∃E twice, with fresh names]. Then $p^2 = r^2q^2 = 2q^2$ [arith], so p^2 is even [arith], and by Theorem 2.1.3 p is even [∀E at p, →E]. Writing $p = 2m$, we get $4m^2 = 2q^2$, so $q^2 = 2m^2$ is even [arith], and by the same theorem q is even [∀E at q, →E]. So p and q are both even [∧I], while they were chosen not both even: a contradiction [¬E; then ∃E twice, since ⊥ mentions neither p nor q]. Hence $r^2 \neq 2$ [¬I]. Since r was arbitrary, no rational number has square 2 [∀I].

Level 1. Level 1 keeps every box, every assumption and every closing line, and compresses the work between them into three lines.

1	$\forall r \in \mathbb{Q}\ \exists p \in \mathbb{Z}\ \exists q \in \mathbb{Z}\ L(r,p,q)$	
2	r_0 | $r_0 \in \mathbb{Q}$	asm, for ∀I
3	$r_0 \cdot r_0 = 2$	asm, for ¬I
4	$\exists p \in \mathbb{Z}\ \exists q \in \mathbb{Z}\ L(r_0,p,q)$	∀E 1, 2
5	p_0 | $p_0 \in \mathbb{Z}$	asm, for ∃E
6	$\exists q \in \mathbb{Z}\ L(r_0,p_0,q)$	asm, for ∃E
7	q_0 | $q_0 \in \mathbb{Z}$	asm, for ∃E
8	$L(r_0,p_0,q_0)$	asm, for ∃E
9	$E(p_0)$	∧E; arith; lemma 3, 5, 7, 8
10	$E(q_0)$	arith; lemma 3, 5, 7, 8, 9
11	⊥	∧E, ∧I, ¬E 8, 9, 10
12	⊥	∃E 6, 7–11
13	⊥	∃E 4, 5–12
14	$\neg(r_0 \cdot r_0 = 2)$	¬I 3–13
15	$\forall r \in \mathbb{Q}\ \neg(r \cdot r = 2)$	∀I 2–14

Line 9 takes the first conjunct out of $L(r_0,p_0,q_0)$, and line 11 the second.

The answer to the prompt: "By (L), write $r = p/q$ with integers p, q, not both even" stands for line 4 (∀E of premise 1 at r_0) and for the openings of the two ∃E boxes, lines 5–8. It also displays the conjunct $r = p/q$ on its own, so the ∧E inside line 9 that takes it out has nothing left to say.

↻ **Second reading** · *Level 0 of the square root of two*

Level 0. Premise 1 is (L) in the form above, premise 2 is Theorem 2.1.3. The derivation has 24 lines.

1		$\forall r \in \mathbb{Q}\ \exists p \in \mathbb{Z}\ \exists q \in \mathbb{Z}\ (p = r \cdot q \wedge \neg(E(p) \wedge E(q)))$	
2		$\forall n \in \mathbb{Z}\ (E(n \cdot n) \to E(n))$	
3	r_0	$r_0 \in \mathbb{Q}$	asm, for $\forall$I
4		$r_0 \cdot r_0 = 2$	asm, for $\neg$I
5		$\exists p \in \mathbb{Z}\ \exists q \in \mathbb{Z}\ (p = r_0 \cdot q \wedge \neg(E(p) \wedge E(q)))$	$\forall$E 1, 3
6	p_0	$p_0 \in \mathbb{Z}$	asm, for $\exists$E
7		$\exists q \in \mathbb{Z}\ (p_0 = r_0 \cdot q \wedge \neg(E(p_0) \wedge E(q)))$	asm, for $\exists$E
8	q_0	$q_0 \in \mathbb{Z}$	asm, for $\exists$E
9		$p_0 = r_0 \cdot q_0 \wedge \neg(E(p_0) \wedge E(q_0))$	asm, for $\exists$E
10		$p_0 = r_0 \cdot q_0$	$\wedge$E 9
11		$p_0 \cdot p_0 = 2 \cdot (q_0 \cdot q_0)$	arith 4, 10
12		$E(p_0 \cdot p_0)$	arith 8, 11
13		$E(p_0 \cdot p_0) \to E(p_0)$	$\forall$E 2, 6
14		$E(p_0)$	$\to$E 13, 12
15		$E(q_0 \cdot q_0)$	arith 6, 8, 11, 14
16		$E(q_0 \cdot q_0) \to E(q_0)$	$\forall$E 2, 8
17		$E(q_0)$	$\to$E 16, 15
18		$E(p_0) \wedge E(q_0)$	$\wedge$I 14, 17
19		$\neg(E(p_0) \wedge E(q_0))$	$\wedge$E 9
20		$\bot$	$\neg$E 18, 19
21		$\bot$	$\exists$E 7, 8–20
22		$\bot$	$\exists$E 5, 6–21
23		$\neg(r_0 \cdot r_0 = 2)$	$\neg$I 4–22
24		$\forall r \in \mathbb{Q}\ \neg(r \cdot r = 2)$	$\forall$I 3–23

The three "arith" lines are facts about integers, each checkable by hand. Line 11: $p_0 \cdot p_0 = (r_0 q_0)(r_0 q_0) = (r_0 \cdot r_0)(q_0 \cdot q_0) = 2 \cdot (q_0 \cdot q_0)$, by lines 10 and 4. Line 12: $p_0 \cdot p_0$ is 2 times the integer $q_0 \cdot q_0$, so it is even. Line 15: by line 14, $p_0 = 2m$ for an integer m; then line 11 reads $4m^2 = 2q_0^2$, so $q_0^2 = 2m^2$ is even. The side conditions at the flagged lines hold: at line 21 the flag q_0 is not in $\bot$, in line 7, in a premise, or in the assumptions still open (lines 3, 4, 6, 7); at line 22 the flag p_0 is not in $\bot$, in line 5, in a premise, or in lines 3 and 4; at line 24 the flag r_0 is in no premise and not in line 24.

Level-1 line 9 is level-0 lines 10–14; line 10 is lines 15–17; line 11 is lines 18–20. Premise 2 is not a line of level 1; it is cited as "lemma". Every other line of level 1 is a line of level 0, renumbered.

At level 0, "By (L), write $r = p/q$ with integers p, q, not both even" stands for line 5 ($\forall$E of premise 1 at r_0) and for the openings of the two $\exists$E boxes, lines 6–9. It also displays the conjunct, in the form $r = p/q$, on its own, so the $\wedge$E of line 10 that extracts it has nothing left

to say and can be silent.

A check at level 0.

Quick check. Which lines of the level-0 derivation in Example (The irrationality of the square root of two at four levels) have no counterpart in the level-3 text, and which item of Definition 8.3.2 lets each go?

Solution. Eight lines. Line 5 is ∀E of premise 1 at the obvious term r_0, the name in sight; "By (L)" cites premise 1, and the instance is not written. Lines 10 and 19 are ∧E. Lines 13 and 16 are ∀E of the lemma at the obvious terms p_0 and q_0; the citations "by Theorem 2.1.3" and "by the same theorem" remain, and they are the →E lines 14 and 17. Lines 21 and 22 are the closing lines of the two ∃E boxes. Line 24 is the closing line of ∀I: level 3 never says "since r was arbitrary". Every other line has a phrase: lines 1 and 2, the premises, are the citations "By (L)" and "by Theorem 2.1.3"; lines 3, 4, 6–9 the openings; lines 11, 12, 15 the arithmetic; lines 14, 17 the citations; lines 18 and 20 the clash, "both even, a contradiction", which names the pair through "not both even" two sentences earlier; and line 23, "hence $r^2 \neq 2$", the conclusion of ¬I, which is never silent.

Look at what the four texts share. Every level has the box of r, the box of the assumption $r^2 = 2$, the choice of p and q, two uses of the lemma, and a contradiction followed by "hence $r^2 \neq 2$". That common part is the proof; the rest is how much of the routine work each reader is trusted to do. The skeleton is the place where you can see it most clearly: fifteen lines, of which only lines 9 and 10 carry the mathematics; every other line is a decision about logic.

Quick check. Which lines of the level-1 skeleton in Example (The irrationality of the square root of two at four levels) have no counterpart in the level-3 text, and which item of Definition 8.3.2 lets each go?

Solution. Four lines. Line 4 is ∀E of premise 1 at the obvious term r_0, the name in sight; "By (L)" cites premise 1, and the instance is not written. Lines 12 and 13 are the closing lines of the two ∃E boxes. Line 15 is the closing line of ∀I: level 3 never says "since r was arbitrary". Every other line has a phrase: line 1 is "By (L)"; lines 2, 3 and 5–8 the openings; lines 9 and 10 the computations with their citations of Theorem 2.1.3; line 11 the clash, "both even, a contradiction"; and line 14, "hence $r^2 \neq 2$", the conclusion of ¬I, which is never silent.

8.3.4 Uniform continuity, twice

The second example has no arithmetic at all. That makes it a good test of the levels: what is left when there is nothing to compress?

Example. Uniformly continuous implies continuous, at four levels

Write the proof of Example (Uniformly continuous implies continuous), that a function uniformly continuous on A is continuous at every point of A, at levels 1 and 3, and say where each level records what δ depends on.

Solution. **Level 1.** Every line of the full derivation of Example (Uniformly continuous implies continuous) other than the premise opens or closes a box, or is cited by a closing line, except one stretch inside the box of δ_0, which becomes line 7.

1		$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A\ \forall y \in A\ (\lvert x - y \rvert < \delta \to \lvert f(x) - f(y) \rvert < \varepsilon)$	
2	x_0	$x_0 \in A$	asm, for $\forall$I
3	ε_0	$\varepsilon_0 > 0$	asm, for $\forall$I
4		$\exists \delta > 0\ \forall x \in A\ \forall y \in A\ (\lvert x - y \rvert < \delta \to \lvert f(x) - f(y) \rvert < \varepsilon_0)$	$\forall$E 1, 3
5	δ_0	$\delta_0 > 0$	asm, for $\exists$E
6		$\forall x \in A\ \forall y \in A\ (\lvert x - y \rvert < \delta_0 \to \lvert f(x) - f(y) \rvert < \varepsilon_0)$	asm, for $\exists$E
7		$\exists \delta > 0\ \forall y \in A\ (\lvert x_0 - y \rvert < \delta \to \lvert f(x_0) - f(y) \rvert < \varepsilon_0)$	$\forall$E, $\exists$I 2, 5, 6
8		$\exists \delta > 0\ \forall y \in A\ (\lvert x_0 - y \rvert < \delta \to \lvert f(x_0) - f(y) \rvert < \varepsilon_0)$	$\exists$E 4, 5–7
9		$\forall \varepsilon > 0\ \exists \delta > 0\ \forall y \in A\ (\lvert x_0 - y \rvert < \delta \to \lvert f(x_0) - f(y) \rvert < \varepsilon)$	$\forall$I 3–8
10		$\forall x \in A\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall y \in A\ (\lvert x - y \rvert < \delta \to \lvert f(x) - f(y) \rvert < \varepsilon)$	$\forall$I 2–9

Level 3.

> Let $x \in A$ and $\varepsilon > 0$. By uniform continuity, choose $\delta > 0$ (depending on ε only) such that $\lvert f(u) - f(v) \rvert < \varepsilon$ for all $u, v \in A$ with $\lvert u - v \rvert < \delta$. In particular $\lvert f(x) - f(y) \rvert < \varepsilon$ for every $y \in A$ with $\lvert x - y \rvert < \delta$, so this δ works for x and ε.

Where the dependence of δ is recorded. At level 1, in two places: the box of δ_0 lies inside the boxes of x_0 and ε_0, so δ_0 **may** depend on both; and the line it comes from, line 4, mentions ε_0 and not x_0, so it **does** depend only on ε_0 (Section 6.4). At level 3 it is the parenthesis "(depending on ε only)", and nothing else: the parenthesis is the whole of the box structure that survives. Section 8.2's picture draws the same record.

↻ **Second reading** · *Levels 0 and 2 of the same proof*

Level 0 is the derivation of Example (Uniformly continuous implies continuous), copied here with the same line numbers. Line 1 is uniform continuity (UC), line 11 is continuity at every point (C).

1	$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x \in A\ \forall y \in A\ (\lvert x - y \rvert < \delta \to \lvert f(x) - f(y) \rvert < \varepsilon)$	
2	x_0 | $x_0 \in A$	asm, for ∀I
3	ε_0 | $\varepsilon_0 > 0$	asm, for ∀I
4	$\exists \delta > 0\ \forall x \in A\ \forall y \in A\ (\lvert x - y \rvert < \delta \to \lvert f(x) - f(y) \rvert < \varepsilon_0)$	∀E 1, 3
5	δ_0 | $\delta_0 > 0$	asm, for ∃E
6	$\forall x \in A\ \forall y \in A\ (\lvert x - y \rvert < \delta_0 \to \lvert f(x) - f(y) \rvert < \varepsilon_0)$	asm, for ∃E
7	$\forall y \in A\ (\lvert x_0 - y \rvert < \delta_0 \to \lvert f(x_0) - f(y) \rvert < \varepsilon_0)$	∀E 6, 2
8	$\exists \delta > 0\ \forall y \in A\ (\lvert x_0 - y \rvert < \delta \to \lvert f(x_0) - f(y) \rvert < \varepsilon_0)$	∃I 5, 7
9	$\exists \delta > 0\ \forall y \in A\ (\lvert x_0 - y \rvert < \delta \to \lvert f(x_0) - f(y) \rvert < \varepsilon_0)$	∃E 4, 5–8
10	$\forall \varepsilon > 0\ \exists \delta > 0\ \forall y \in A\ (\lvert x_0 - y \rvert < \delta \to \lvert f(x_0) - f(y) \rvert < \varepsilon)$	∀I 3–9
11	$\forall x \in A\ \forall \varepsilon > 0\ \exists \delta > 0\ \forall y \in A\ (\lvert x - y \rvert < \delta \to \lvert f(x) - f(y) \rvert < \varepsilon)$	∀I 2–10

Level 1. Every line of level 0 other than the premise opens or closes a box, or is cited by a closing line, except line 7: lines 7 and 8, which specialize the chosen δ_0 to the point x_0 and repackage the result, are the one stretch of work inside a box.

Level 2.

> Let $x \in A$ [∀I] and let $\varepsilon > 0$ [∀I]. By uniform continuity, applied at ε [∀E], there is $\delta > 0$ such that $\lvert u - v \rvert < \delta$ implies $\lvert f(u) - f(v) \rvert < \varepsilon$ for all $u, v \in A$; choose such a δ, which depends on ε only [∃E]. Taking $u = x$ [∀E], every $y \in A$ with $\lvert x - y \rvert < \delta$ has $\lvert f(x) - f(y) \rvert < \varepsilon$. So this δ works for x and ε [∃I; then ∃E closes, since "some δ works" does not mention the chosen one]. Since ε and x were arbitrary, f is continuous at every point of A [∀I twice].

As in Section 6.4, the prose renames the bound variables of UC to u, v, so that x can name the fixed point.

Where the dependence of δ is recorded. At level 0, in two places: the box of δ_0 lies inside the boxes of x_0 and ε_0, so δ_0 **may** depend on both; and the line it comes from, line 4, mentions ε_0 and not x_0, so it **does** depend only on ε_0 (Section 6.4). At level 1 the record is the same, lines 3 to 5 being unchanged. At level 2 it is the clause "which depends on ε only", placed at the ∃E tag, after the ∀E at ε that produced it. At level 3 it is the parenthesis "(depending on ε only)", and nothing else: the parenthesis is the whole of the box structure that survives.

Here the gap between level 0 and level 1 is one line, because there is no arithmetic to compress.

From level 1 to level 3, ten lines become three sentences. What the level-3 text keeps is exactly the never-silent part of Definition 8.3.2: the openings of the two ∀I boxes ("let $x \in A$ and $\varepsilon > 0$"), the citation and the choice with its fresh name ("by uniform continuity, choose δ"), the dependence of the witness, and the witness itself ("this δ works"). The ∀E at $u = x$, inside line 7, is the obvious instance and becomes "in particular". The ∃E closing line 8 and the two ∀I closing lines 9 and 10 are silent.

Quick check. In the level-3 text, which words record that δ was chosen inside the box of ε from a line that mentions ε only?

Solution. The parenthesis "(depending on ε only)".

8.3.5 It suffices to show, as a column

Section 8.2 read "it suffices to show φ" as a step that replaces the goal by φ and shows on the spot that φ is enough. A proof that makes several such steps in a row is easiest to plan, and to check, as a **column**. The goal goes at the top. Each row is a claim from which the row above follows, with the reason beside it. At the bottom is a claim that needs no reduction: a hypothesis, a fact or a cited result. Read upwards, the column is the proof.

Example. Two columns

(a) Prove that $x^2 - 6x + 10 > 0$ for every real x. (b) Let $a_n := (2n+1)/(n+1)$ for $n \in \mathbb{N}$. Prove that $a_n \to 2$. Plan each proof as a column of reductions, then write it.

Solution. (a)

It suffices to show	because
$x^2 - 6x + 10 > 0$	(the goal)
$(x-3)^2 + 1 > 0$	$x^2 - 6x + 10 = (x-3)^2 + 1$
$(x-3)^2 \geq 0$	a number at least 0, plus 1, is positive
—	every square is at least 0

(b) Let $\varepsilon > 0$. For $n \in \mathbb{N}$ with $n \geq N$, where N is still to be chosen:

It suffices to show	because
$\lvert a_n - 2\rvert < \varepsilon$	(the goal, for $n \geq N$)
$1/(n+1) < \varepsilon$	$a_n - 2 = -1/(n+1)$, so $\lvert a_n - 2\rvert = 1/(n+1)$
$n + 1 > 1/\varepsilon$	for positive numbers, taking reciprocals reverses $<$
$N > 1/\varepsilon$	$n \geq N$ gives $n + 1 > N$
—	choose an integer $N > 1/\varepsilon$ (the school fact of Example (The reciprocals tend to zero)); since $1/\varepsilon > 0$, $N > 0$, so $N \in \mathbb{N}$; it depends on ε

Each reason is an implication that points **up**: the row implies the row above it. Read from the bottom, (a) is the proof "Since $(x-3)^2 \geq 0$, $x^2 - 6x + 10 = (x-3)^2 + 1 \geq 1 > 0$." Read the same way, (b) is "Let $\varepsilon > 0$, and choose $N \in \mathbb{N}$ with $N > 1/\varepsilon$. For $n \geq N$, $\lvert a_n - 2\rvert = 1/(n+1) < 1/N < \varepsilon$." In (b) the constant is found at the bottom of the column, last, and written in the proof's first sentence, as step 5 of Section 8.5 will require.

Quick check. In column (b), which row introduces a constant, and which reason needs $\varepsilon > 0$?

Solution. The bottom row introduces N, chosen from the school fact at $1/\varepsilon$, so it depends on ε. The reciprocal row needs $\varepsilon > 0$: $1/\varepsilon$ must exist and be positive for "$n + 1 > 1/\varepsilon$ gives $1/(n+1) < \varepsilon$". The bottom row uses it too: $1/\varepsilon$ must be defined, and an integer above $1/\varepsilon$ lies in $\mathbb{N}$ because $1/\varepsilon > 0$.

8.3.6 Expand to check

The levels are not only a way to write. They are a way to check. When a sentence of a level-3 proof leaves you uneasy, rewrite that stretch, and only that stretch, at level 1. Either every line of the expansion is legal, and the doubt is settled, or one line breaks a rule, and you have found the error and its exact address.

Start with a sentence whose fault is close to the surface:

> Some element of A is positive; let x be one. Since x was arbitrary, every element of A is positive.

Expand it. "Some element of A is positive" is the line $\exists x\,(x \in A \wedge x > 0)$; "let x be one" opens an $\exists$E box with a fresh name x_0; "since x was arbitrary" needs a $\forall$I. At level 1, with that line marked:

1		$\exists x\,(x \in A \wedge x > 0)$	
2	x_0	$x_0 \in A \wedge x_0 > 0$	
3		$x_0 \in A$	asm, for $\to$I
4		$x_0 > 0$	$\wedge$E 2
5		$x_0 \in A \to x_0 > 0$	$\to$I 3–4
6		$\forall x\,(x \in A \to x > 0)$	$\forall$I? 2–5

Line 6 is what "since x was arbitrary" needs, and no rule gives it. It cannot cite the box 2–5, which is still open (and is the $\exists$E box of x_0); and a new $\forall$I box flagged x_0, opened inside it, breaks the side condition of $\forall$I (Definition 5.4.1), because x_0 appears in the open assumption on line 2. That assumption says that x_0 is positive, so x_0 is one particular element, not an arbitrary one. Section 8.6 catalogues the error as (N3). For $A = \{-1, 1\}$ the premise is true and the conclusion false.

Here is a harder one, which sounds like the level-3 proof of Example (Uniformly continuous implies continuous, at four levels) run backwards.

> Let $\varepsilon > 0$. For each $x \in A$, continuity at x gives some $\delta > 0$ such that $|f(x)-f(y)| < \varepsilon$ for every $y \in A$ with $|x - y| < \delta$. Hence, choosing δ as above for every x, we have $|f(x) - f(y)| < \varepsilon$ whenever $x, y \in A$ and $|x - y| < \delta$, so f is uniformly continuous on A.

Expand it. "Let $\varepsilon > 0$" opens a box for ε_0. "For each $x \in A$" opens a box for x_0. "Continuity at x gives some δ" is $\forall$E twice and the opening of an $\exists$E box for δ_0. At level 1, with the line the sentence needs as line 7, its rule marked with a question mark:

1			$\forall x \in A\,\forall \varepsilon > 0\,\exists \delta > 0\,\forall y \in A\,(\lvert x - y\rvert < \delta \to \lvert f(x) - f(y)\rvert < \varepsilon)$	
2	ε_0		$\varepsilon_0 > 0$	
3		x_0	$x_0 \in A$	
4			$\exists \delta > 0\,\forall y \in A\,(\lvert x_0 - y\rvert < \delta \to \lvert f(x_0) - f(y)\rvert < \varepsilon_0)$	$\forall$E twice 1, 3, 2
5		δ_0	$\delta_0 > 0$	
6			$\forall y \in A\,(\lvert x_0 - y\rvert < \delta_0 \to \lvert f(x_0) - f(y)\rvert < \varepsilon_0)$	
7			$\forall x \in A\,\forall y \in A\,(\lvert x - y\rvert < \delta_0 \to \lvert f(x) - f(y)\rvert < \varepsilon_0)$	$\forall$I? 3–6

Line 7 is what the phrase "choosing δ as above for every x" needs: one δ_0 that works for all x and y, so that $\exists$I can then produce UC's $\exists\delta$ inside the box of ε_0. But line 7 would be a $\forall$I over x_0, and no such step is legal here. The box of x_0 opened at line 3 is still open, so $\forall$I cannot cite it; and a new box with the flag x_0, opened inside the box of δ_0, breaks the side condition of $\forall$I (Definition 5.4.1), because x_0 appears in the open assumptions on lines 3 and 6. The phrase "for every x" has generalized over a name on which the chosen δ_0 depends. This is the first blocked line of Example (The search from C to UC), met there in the search; here it was found by expanding one English sentence. Section 8.6 catalogues the error, and Example (The square function on the real line) shows that no repair is possible: for x^2 on $\mathbb{R}$ the conclusion is false.

The expansion took six lines, and it was enough to expand the one doubtful stretch. That is

the general pattern: you never need to rewrite a whole proof at level 0 to check it. Find the sentence you do not trust, find the boxes open at that sentence, and write the few lines it stands for.

Remark (Looking ahead). From *Analysis on the Real Line*, Chapter 1 on, proofs are written at level 3, and a rule tag appears only where the rule is itself the point; *Sets, Relations, and Cardinality*, Chapters 1–3 may still tag some proofs at level 2.

8.3.7 Exercises

8.3.7.1 A. Check your understanding

Exercise. A1

Parts (a)–(c) are about Example (Uniformly continuous implies continuous, at four levels), parts (d) and (e) about Example (The irrationality of the square root of two at four levels).

(a) Which lines of the level-1 skeleton does the level-3 sentence "Let $x \in A$ and $\varepsilon > 0$" stand for? Which open a box, which close one, and which are written at level 3?

(b) Where is the dependence of δ recorded at levels 1 and 3?

(c) State the idea of the proof in two sentences.

(d) At which line of levels 1, 2 and 3 is the "not both even" clause of (L) used to reach the contradiction?

(e) True or false, with a reason: at level 3, the sentence "hence $r^2 \neq 2$" may be left out of the proof of Theorem 2.2.1, since the reader has just been told that a contradiction was reached.

Solution. (a) Lines 2, 3, 9 and 10. Lines 2 and 3 open the flagged boxes of x_0 and ε_0, with their bounds $x_0 \in A$ and $\varepsilon_0 > 0$; lines 9 and 10 close them by $\forall$I. The openings are written ("let"); the closings are silent at level 3 (Definition 8.3.2).

(b) At level 1: the box of δ_0 sits inside the boxes of x_0 and ε_0, and line 4, the line it comes from, mentions ε_0 only. At level 3: the parenthesis "(depending on ε only)".

(c) Uniform continuity at ε provides one δ that serves all pairs of points of A at once, and it provides it before any point is named. Continuity at a point x asks only for the pairs whose first point is x, so the same δ works.

(d) Level 1: it is the second conjunct of $L(r_0, p_0, q_0)$ on line 8, used at line 11. Level 2: "while they were chosen not both even: a contradiction". Level 3: "Thus p and q are both even, a contradiction", where the clash refers back to "not both even" in the sentence that chose p and q.

(e) False. "Hence $r^2 \neq 2$" is the conclusion of $\neg$I, which Definition 8.3.2 puts among the steps that are never silent. A contradiction alone does not say which assumption it refutes; in this proof the boxes of p, q and of $r^2 = 2$ are all open when it is reached, and the sentence says which box closes and what it gives.

Exercise. A2: Drill

Each part takes a line or two. (a) names a level; (b)–(f) decide whether a step may be silent at level 3; (g) and (h) judge a text; (i) and (j) decide whether a line may be justified TAUT; (k) and (l) work with an "it suffices to show" column.

(a) "Let $\varepsilon > 0$ [$\forall$I]. Take $\delta := \varepsilon/2$ [$\exists$I]. . . . " At which level is this text written?

(b) At level 3, may "since x was arbitrary" be left out?

(c) May "choose k with $n = 2k$" be left out?

(d) May "by the triangle inequality" be left out?

(e) May "in particular $x > 0$", after "since $0 < x < 1$", be left out?

(f) May the split "$x \geq 0$ or $x < 0$, and every real x is one or the other" be left out?

(g) A skeleton has the line "$|f(x) - f(y)| < \varepsilon$, arith". Is the proof finished?

(h) A text says "Let r be rational. Then p is even." Name two never-silent steps that are missing.

(i) From Section 6.3 on, may the level-1 line $x_0 > 0$, citing the line $x_0 > 0 \wedge y_0 > 0$, be justified TAUT?

(j) May the line $\forall y \forall x\, R(x, y)$, citing $\forall x \forall y\, R(x, y)$, be justified TAUT?

(k) A column for "$2^{n+1} > n + 1$, given $2^n > n$ and $n \geq 1$" has the first row "it suffices to show $2^n + 2^n > n + 1$". What is that row's reason, and what row could come next?

(l) In an "it suffices to show" column, which row is proved outright, with no further reduction?

Solution.

(a) Level 2: sentences, each tagged with the rule it carries.

(b) Yes: it is the closing of $\forall$I.

(c) No: it is an $\exists$E with a fresh name.

(d) No: it is a citation.

(e) Yes: it is $\wedge$E.

(f) No: it is the split of $\vee$E, with the reason it exhausts.

(g) No: a skeleton is a plan, and every arith line must be expanded (the warning "A skeleton is not a finished proof").

(h) The choice of p ($\exists$E with a fresh name, citing (L)): p appears from nowhere. Also missing: the assumption $r^2 = 2$, the opening of $\neg$I, and the citation of Theorem 2.1.3 that licenses "p is even".

(i) Yes. On the skeleton (Proposition 6.3.1) it is $P \wedge Q \models P$, true in every row: it is $\wedge$E.

(j) No. On the skeleton the two lines are different letters, and $P \models Q$ fails in a row (P true, Q false). The step needs $\forall$E twice and $\forall$I twice.

(k) $2^{n+1} = 2 \cdot 2^n = 2^n + 2^n$ (=E). Next: "it suffices to show $2^n \geq 1$", because then $2^n + 2^n > n + 2^n \geq n + 1$ by the hypothesis $2^n > n$.

(l) The bottom row. It is a hypothesis, a fact or a cited result; every other row is reduced to the one below it.

8.3.7.2 B. Practice

Exercise. B1: Some-all gives all-some, on the natural numbers

Read Proposition 5.6.2 over the domain $\mathbb{N}$, with $R(x, y)$ as "$x \leq y$": from $\exists x\, \forall y\, (x \leq y)$, conclude $\forall y\, \exists x\, (x \leq y)$. The premise is true, with the witness $x = 0$. Write the proof at levels 1, 2 and 3. Your proof must use the premise, although this instance of the conclusion also has a one-line proof, with $x = y$.

Solution. **Level 1.** The proof has no arithmetic, so the skeleton is the level-0 derivation of Proposition 5.6.2 with $R(x, y)$ read as $x \leq y$:

1	$\exists x\, \forall y\, (x \leq y)$		
2	y_0 | x_0 | $\forall y\, (x_0 \leq y)$		asm, for $\exists$E
3		$x_0 \leq y_0$	$\forall$E 2
4		$\exists x\, (x \leq y_0)$	$\exists$I 3
5		$\exists x\, (x \leq y_0)$	$\exists$E 1, 2–4
6	$\forall y\, \exists x\, (x \leq y)$		$\forall$I 2–5

Level 2.

> Let $y \in \mathbb{N}$ [$\forall$I]. By hypothesis, some natural number is at most every natural number; choose such a number x_0, which depends on nothing, since the hypothesis mentions no y [$\exists$E, x_0 fresh]. In particular $x_0 \leq y$ [$\forall$E at y], so some natural number is at most y [$\exists$I with witness x_0; then $\exists$E closes, since this does not mention x_0]. Since y was arbitrary, every natural number has a natural number at most it [$\forall$I].

Level 3.

> Let $y \in \mathbb{N}$. By hypothesis, choose $x_0 \in \mathbb{N}$ with $x_0 \leq z$ for every $z \in \mathbb{N}$; the hypothesis does not mention y, so x_0 does not depend on y. In particular $x_0 \leq y$, so x_0 is a natural number at most y.

The level-3 text renames the bound variable of the hypothesis to z, so that y can name the fixed number. The witness x_0 was chosen inside the box of y_0, but it comes from line 1, which mentions no y; so it does not depend on y. The one-line proof, "$y \leq y$, so take $x = y$", is $\exists$I with the witness y, which **does** depend on y. Both prove the conclusion; only the first is an instance of Proposition 5.6.2, and only the first shows that a single x serves every y.

Exercise. B2: Back from prose to a skeleton, with blanks

Here is a level-3 proof.

> **Claim.** If n is an odd integer, then $8 \mid n^2 - 1$.
>
> Let n be odd, say $n = 2k + 1$ with k an integer. Then $n^2 - 1 = 4k(k + 1)$, so it suffices to show that $k(k + 1)$ is even. If $k = 2j$ for an integer j, then $k(k + 1) = 2 \cdot j(2j + 1)$; if $k = 2j + 1$, then $k(k + 1) = 2 \cdot (2j + 1)(j + 1)$. Either

way $k(k+1) = 2i$ for an integer i, and then $n^2 - 1 = 8i$.

Work over the domain $\mathbb{Z}$, with "n is odd" as $\exists k\,(n = 2k+1)$ and $8 \mid m$ as $\exists t\,(m = 8t)$. Here is its level-1 skeleton with nine blanks, labeled (i)–(ix). *Name the move, then complete it*, and say where the proof uses the fact (E1) of Section 2.1, that every integer is even or odd.

1	n_0	$\exists k\,(n_0 = 2 \cdot k + 1)$	
2	k_0	$n_0 = 2 \cdot k_0 + 1$	
3		$n_0 \cdot n_0 - 1 = 4 \cdot (k_0 \cdot (k_0 + 1))$	arith 2
4		(i)	(E1)
5		$\exists j\,(k_0 = 2 \cdot j)$	
6		j_0 $k_0 = 2 \cdot j_0$	
7		$k_0 \cdot (k_0 + 1) = 2 \cdot (j_0 \cdot (2 \cdot j_0 + 1))$	arith 6
8		$\exists i\,(k_0 \cdot (k_0 + 1) = 2 \cdot i)$	∃I 7
9		$\exists i\,(k_0 \cdot (k_0 + 1) = 2 \cdot i)$	(ii)
10		$\exists j\,(k_0 = 2 \cdot j + 1)$	
11		j_1 $k_0 = 2 \cdot j_1 + 1$	
12		$k_0 \cdot (k_0 + 1) = 2 \cdot ((2 \cdot j_1 + 1) \cdot (j_1 + 1))$	arith 11
13		$\exists i\,(k_0 \cdot (k_0 + 1) = 2 \cdot i)$	∃I 12
14		$\exists i\,(k_0 \cdot (k_0 + 1) = 2 \cdot i)$	(iii)
15		$\exists i\,(k_0 \cdot (k_0 + 1) = 2 \cdot i)$	(iv)
16		i_0 (v)	
17		$n_0 \cdot n_0 - 1 = 8 \cdot i_0$	arith 3, 16
18		$\exists t\,(n_0 \cdot n_0 - 1 = 8 \cdot t)$	∃I 17
19		$\exists t\,(n_0 \cdot n_0 - 1 = 8 \cdot t)$	(vi)
20		$\exists t\,(n_0 \cdot n_0 - 1 = 8 \cdot t)$	(vii)
21		$\exists k\,(n_0 = 2 \cdot k + 1) \to \exists t\,(n_0 \cdot n_0 - 1 = 8 \cdot t)$	(viii)
22		$\forall n\,(\exists k\,(n = 2 \cdot k + 1) \to \exists t\,(n \cdot n - 1 = 8 \cdot t))$	(ix)

Solution. The move is skeleton first: every box closes by the rule its opening asks for: a "choose" by ∃E on a line that does not mention the chosen name, the split by ∨E, "suppose n is odd" by →I, "let n" by ∀I.

Completion: (i) $\exists j\,(k_0 = 2 \cdot j) \vee \exists j\,(k_0 = 2 \cdot j + 1)$, by (E1) at k_0; (ii) ∃E 5, 6–8; (iii) ∃E 10, 11–13; (iv) ∨E 4, 5–9, 10–14; (v) $k_0 \cdot (k_0 + 1) = 2 \cdot i_0$, with i_0 fresh; (vi) ∃E 15, 16–18; (vii) ∃E 1, 2–19; (viii) →I 1–20; (ix) ∀I 1–21.

(E1) enters at line 4; in the prose it is "if $k = 2j$ … if $k = 2j+1$", whose reason for covering every case is not written, so the given proof is slightly short of level 3 there.

Exercise. B3: A column of reductions

(a) Prove that $|x-1|+|x+1| \geq 2$ for every real x by an "it suffices to show" column, each row with its reason.

(b) Turn the column into a level-3 proof.

(c) Which reasons in your column are =E steps that level 3 may leave silent?

(d) A student's column has the row "it suffices to show $|x-1| \geq 1$ and $|x+1| \geq 1$". Is that row's reason correct? Can the student finish?

Solution. (a)

It suffices to show	because
$\lvert x-1\rvert+\lvert x+1\rvert \geq 2$	(the goal)
$\lvert 1-x\rvert+\lvert x+1\rvert \geq \lvert(1-x)+(x+1)\rvert$	$\lvert x-1\rvert=\lvert 1-x\rvert$, and $(1-x)+(x+1)=2$, so the right side is $\lvert 2\rvert=2$
—	the triangle inequality, Example (The triangle inequality by cases) with $1-x$ and $x+1$ in place of its x and y

(b) "By the triangle inequality, $|x-1|+|x+1| = |1-x|+|x+1| \geq |(1-x)+(x+1)| = 2$."

(c) $|x-1| = |1-x|$ and $(1-x)+(x+1) = 2$: both are =E links of a chain, silent at level 3 (Definition 8.3.2). The citation of the triangle inequality is never silent.

(d) The reason is correct: if both hold, the sum is at least 2, so the row implies the goal. But the student cannot finish. The new row is false at $x = \frac{1}{2}$, where $|x-1| = \frac{1}{2}$, although the goal holds there ($\frac{1}{2}+\frac{3}{2}=2$). A reduction is legal whenever its arrow points up; it is useful only when the new row is true.

Marking scheme. (10 points) - 3: (a) a column whose every row has an upward-pointing reason, ending in a cited result or a fact (the triangle inequality, or a case split on the sign of $x-1$ and $x+1$). - 3: (b) a forward level-3 text with the citation written. - 2: (c) the two =E links named as silent, the citation as spoken. - 2: (d) the reason judged correct, and the row shown false at a named x: it fails exactly when $0 < |x| < 2$, for instance at $x = \frac{1}{2}$.

A three-case proof ($x \leq -1$; $-1 < x < 1$; $x \geq 1$) written as a column earns (a) and (b) in full. Fatal errors (0 for the step): a row whose reason points down ("the goal implies the row"); a case split that does not exhaust (N8).

8.3.7.3 C. Going deeper

Exercise. C1: Arbitrarily close numbers, at two levels

Consider Proposition 6.2.5: for real numbers a and b, if $|a-b| < \varepsilon$ for every $\varepsilon > 0$, then $a = b$.

(a) Write its proof at level 1, as a skeleton in which school facts about $|\cdot|$ and $<$ are lines justified "arith".

(b) Write it at level 3.

(c) Which lines of your skeleton are silent at level 3? Which instance of ∀E must the level-3 text name, and why?

Solution. *(a)* The statement is an implication, so the outer box is →I; the goal $a = b$ is atomic, and the proof in Section 6.2 opens an RAA box on $a \neq b$.

1	$\forall \varepsilon > 0\ (\lvert a-b\rvert < \varepsilon)$	asm, for →I
2	$\neg(a = b)$	asm, for RAA
3	$\lvert a-b\rvert > 0$	arith 2
4	$\lvert a-b\rvert < \lvert a-b\rvert$	∀E 1, 3
5	$\neg(\lvert a-b\rvert < \lvert a-b\rvert)$	arith
6	$\bot$	¬E 4, 5
7	$a = b$	RAA 2–6
8	$\forall \varepsilon > 0\ (\lvert a-b\rvert < \varepsilon) \to a = b$	→I 1–7

Line 4 is the bounded ∀E at the term $|a-b|$, whose bound $|a-b| > 0$ is line 3. Line 5 is the school fact that no real number is less than itself.

(b)

> Suppose $|a-b| < \varepsilon$ for every $\varepsilon > 0$, and suppose, for a contradiction, that $a \neq b$. Then $|a-b| > 0$, so the hypothesis applies with $\varepsilon = |a-b|$ and gives $|a-b| < |a-b|$; but no real number is less than itself. Hence $a = b$.

(c) Only line 8, the closing line of →I, is silent: "hence if ... then $a = b$" is not written. Line 5 is not silent: it is one member of the clashing pair of the ¬E on line 6, and level 3 names it ("no real number is less than itself"). The instance of line 4 must be named, $\varepsilon = |a-b|$, because it is not the obvious term: no name in sight is the obvious candidate for ε, and the whole proof turns on choosing this one. Saying only "by the hypothesis" would leave the reader to find the instance, and that instance is the idea of the proof. The RAA opening, "suppose, for a contradiction, that $a \neq b$", and its conclusion, "hence $a = b$", are also never silent.

Marking scheme. (10 points) - 2: (a) the →I box on the hypothesis, and inside it a box opened on $a \neq b$. - 2: (a) the box closed by RAA (¬I alone would give only $\neg\neg(a = b)$), with $\bot$ as its last line. - 2: (a) the ∀E at $|a-b|$, citing a line $|a-b| > 0$. - 2: (b) a level-3 text with the opening of the contradiction, the instance, the clash and "hence $a = b$", and no tags. - 2: (c) the silent line identified, and the instance $\varepsilon = |a-b|$ said to be named because it is not the obvious term. - Other routes: a ¬I box on $a \neq b$ followed by ¬¬E, or a split on $|a-b| > 0$ or $|a-b| = 0$ (arith) whose first case reaches $\bot$ by the same ∀E and clash and then $a = b$ by $\bot$E, while the second gives $a = b$ by arith, earns the 4 points of the first two lines, provided the ¬¬E line is written explicitly, or the disjunction of the split is justified (from $|a-b| \geq 0$, or by trichotomy). In the split, the case assumption $|a-b| > 0$ is the line the ∀E cites for the third line, and for (b) and (c) the opening of the ¬I box or of the split stands in for the opening of the contradiction. Fatal errors (0 for the step): the box closed by $\neg I$ with the conclusion $a = b$; the hypothesis applied at an ε not known to be positive.

Exercise. C2: Exit check: levels

Let (a_n) be a sequence of real numbers with $a_n \to L$, and let $c \neq 0$ be a real number. Here is a level-0 derivation that $c\,a_n \to cL$, with the definition of a limit as in Example (Five definitions, translated) (a), the hypothesis $c \neq 0$ as premise 2, and two lines of school arithmetic justified "arith".

1	$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to \lvert a_n - L\rvert < \varepsilon)$	
2	$\neg(c = 0)$	
3	$\varepsilon_0 \mid \varepsilon_0 > 0$	asm, for $\forall$I
4	$\varepsilon_0/\lvert c\rvert > 0$	arith 2, 3
5	$\exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to \lvert a_n - L\rvert < \varepsilon_0/\lvert c\rvert)$	$\forall$E 1, 4
6	$N_0 \mid N_0 \in \mathbb{N}$	asm, for $\exists$E
7	$\forall n \in \mathbb{N}\ (n \geq N_0 \to \lvert a_n - L\rvert < \varepsilon_0/\lvert c\rvert)$	asm, for $\exists$E
8	$n_0 \mid n_0 \in \mathbb{N}$	asm, for $\forall$I
9	$n_0 \geq N_0 \to \lvert a_{n_0} - L\rvert < \varepsilon_0/\lvert c\rvert$	$\forall$E 7, 8
10	$n_0 \geq N_0$	asm, for $\to$I
11	$\lvert a_{n_0} - L\rvert < \varepsilon_0/\lvert c\rvert$	$\to$E 9, 10
12	$\lvert c \cdot a_{n_0} - c \cdot L\rvert < \varepsilon_0$	arith 2, 11
13	$n_0 \geq N_0 \to \lvert c \cdot a_{n_0} - c \cdot L\rvert < \varepsilon_0$	$\to$I 10–12
14	$\forall n \in \mathbb{N}\ (n \geq N_0 \to \lvert c \cdot a_n - c \cdot L\rvert < \varepsilon_0)$	$\forall$I 8–13
15	$\exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to \lvert c \cdot a_n - c \cdot L\rvert < \varepsilon_0)$	$\exists$I 6, 14
16	$\exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to \lvert c \cdot a_n - c \cdot L\rvert < \varepsilon_0)$	$\exists$E 5, 6–15
17	$\forall \varepsilon > 0\ \exists N \in \mathbb{N}\ \forall n \in \mathbb{N}\ (n \geq N \to \lvert c \cdot a_n - c \cdot L\rvert < \varepsilon)$	$\forall$I 3–16

Write the proof at level 2 and at level 3. Then say in one sentence how the case $c = 0$ goes.

Solution. **Level 2.**

Let $\varepsilon > 0$ [$\forall$I]. Since $c \neq 0$, $\lvert c\rvert > 0$, so $\varepsilon/\lvert c\rvert > 0$ [arith]. Applying the definition of $a_n \to L$ with $\varepsilon/\lvert c\rvert$ in place of ε [$\forall$E], choose $N \in \mathbb{N}$, depending on ε (and on the fixed number c), such that $\lvert a_n - L\rvert < \varepsilon/\lvert c\rvert$ for every $n \geq N$ [$\exists$E, N fresh]. Let $n \in \mathbb{N}$ [$\forall$I] with $n \geq N$ [assumption for $\to$I]. Then $\lvert a_n - L\rvert < \varepsilon/\lvert c\rvert$ [$\forall$E, $\to$E], so

$$\lvert c\,a_n - cL\rvert = \lvert c\rvert\,\lvert a_n - L\rvert < \lvert c\rvert \cdot \frac{\varepsilon}{\lvert c\rvert} = \varepsilon$$

[arith]. So this N works for ε [$\to$I, $\forall$I, $\exists$I with witness N; then $\exists$E closes]. Since $\varepsilon > 0$ was arbitrary, $c\,a_n \to cL$ [$\forall$I].

Level 3.

Let $\varepsilon > 0$. Since $c \neq 0$, we have $\varepsilon/\lvert c\rvert > 0$, so by the definition of $a_n \to L$, applied

with $\varepsilon/|c|$, there is $N \in \mathbb{N}$ (depending on ε and c) such that $|a_n - L| < \varepsilon/|c|$ for every $n \geq N$. For every $n \geq N$,

$$|c\,a_n - cL| = |c|\,|a_n - L| < |c| \cdot \frac{\varepsilon}{|c|} = \varepsilon.$$

So this N works for ε, and $c\,a_n \to cL$.

The case $c = 0$. Then $c\,a_n = 0 = cL$ for every n, so $|c\,a_n - cL| = 0 < \varepsilon$ for every $\varepsilon > 0$ and every n, and any N works, for instance $N = 0$; the hypothesis $a_n \to L$ is not needed.

At level 3, the silent steps are those Definition 8.3.2 allows: the ∀E and →E of lines 9 and 11 at the obvious n (the implication of line 11 is line 9, an instance of the property of the chosen N_0, not of a hypothesis or a result), carried by "for every $n \geq N$", and the closing lines 13, 14, 16 and 17. The ∀E of line 5 is **not** silent, because its instance $\varepsilon/|c|$ is not the obvious one; it is named ("applied with $\varepsilon/|c|$"). The witness of line 15 is named ("this N"), and its dependence is stated.

Marking scheme. (10 points) - 4: every never-silent step is present at level 3, one point each: $\varepsilon > 0$ fixed first (∀I); the hypothesis applied at $\varepsilon/|c|$, with its positivity justified, and N chosen from it (∃E); the witness N produced with its dependence on ε stated (∃I); $n \geq N$ taken arbitrary (∀I). - 3: the level-2 tags are correct: ∀I for ε and for n, ∀E at $\varepsilon/|c|$, ∃E for N, →I for $n \geq N$, ∃I for the witness. Deduct one point per wrong or missing tag, down to 0. - 2: no step silent at level 3 beyond those Definition 8.3.2 allows, steps already scored in the first line not counted again; in particular the estimate $|c\,a_n - cL| = |c|\,|a_n - L| < |c| \cdot \varepsilon/|c| = \varepsilon$ is written out for an arbitrary $n \geq N$. - 1: the case $c = 0$ handled, or excluded explicitly with a reason. A level 3 that writes "for every $n \geq N$" as the single phrase for ∀I together with →I earns the ∀I point, as in the solution. Fatal errors (0 for the step): N chosen before ε is fixed; dividing by $|c|$ in the case $c = 0$.

Exercise. C3: Back from prose to a skeleton

Here is a level-3 proof.

> **Claim.** For every integer n, $4 \mid n^2$ or $4 \mid n^2 - 1$.
>
> Let n be an integer. If n is even, say $n = 2k$, then $n^2 = 4k^2$, so $4 \mid n^2$. If n is odd, say $n = 2k + 1$, then $n^2 - 1 = 4k^2 + 4k = 4(k^2 + k)$, so $4 \mid n^2 - 1$.

Reconstruct its level-1 skeleton. Work over the domain $\mathbb{Z}$, with "n is even" as $\exists k\,(n = 2k)$, "n is odd" as $\exists k\,(n = 2k + 1)$ and $4 \mid m$ as $\exists t\,(m = 4t)$. Name the rule that closes each box, and name every step of the skeleton that the prose leaves unsaid.

Hint: The claim is a disjunction for every n. Which disjunct is reached in each case, and by which rule?

Solution. Write $D(n_0)$ for the goal's instance $\exists t\,(n_0 \cdot n_0 = 4 \cdot t) \vee \exists t\,(n_0 \cdot n_0 - 1 = 4 \cdot t)$; in the derivation it is written out.

1	n_0	$\exists k\,(n_0 = 2\cdot k) \vee \exists k\,(n_0 = 2\cdot k+1)$	(E1)
2		$\exists k\,(n_0 = 2\cdot k)$	asm, for $\vee$E
3		k_0 | $n_0 = 2\cdot k_0$	asm, for $\exists$E
4		$n_0\cdot n_0 = 4\cdot(k_0\cdot k_0)$	arith 3
5		$\exists t\,(n_0\cdot n_0 = 4\cdot t)$	$\exists$I 4
6		$\exists t\,(n_0\cdot n_0 = 4\cdot t) \vee \exists t\,(n_0\cdot n_0 - 1 = 4\cdot t)$	$\vee$I 5
7		$\exists t\,(n_0\cdot n_0 = 4\cdot t) \vee \exists t\,(n_0\cdot n_0 - 1 = 4\cdot t)$	$\exists$E 2, 3–6
8		$\exists k\,(n_0 = 2\cdot k+1)$	asm, for $\vee$E
9		k_1 | $n_0 = 2\cdot k_1+1$	asm, for $\exists$E
10		$n_0\cdot n_0 - 1 = 4\cdot(k_1\cdot k_1 + k_1)$	arith 9
11		$\exists t\,(n_0\cdot n_0 - 1 = 4\cdot t)$	$\exists$I 10
12		$\exists t\,(n_0\cdot n_0 = 4\cdot t) \vee \exists t\,(n_0\cdot n_0 - 1 = 4\cdot t)$	$\vee$I 11
13		$\exists t\,(n_0\cdot n_0 = 4\cdot t) \vee \exists t\,(n_0\cdot n_0 - 1 = 4\cdot t)$	$\exists$E 8, 9–12
14		$\exists t\,(n_0\cdot n_0 = 4\cdot t) \vee \exists t\,(n_0\cdot n_0 - 1 = 4\cdot t)$	$\vee$E 1, 2–7, 8–13
15		$\forall n\,(\exists t\,(n\cdot n = 4\cdot t) \vee \exists t\,(n\cdot n - 1 = 4\cdot t))$	$\forall$I 1–14

Line 1 is (E1) at n_0, cited as in B2's line 4. The boxes close by $\exists$E (lines 7 and 13: $D(n_0)$ does not mention k_0 or k_1), $\vee$E (line 14) and $\forall$I (line 15). The second case box could reuse the name k_0, since the two boxes are siblings; k_1 keeps the lines apart. The arithmetic: $(2k_0)^2 = 4k_0^2$ (line 4) and $(2k_1+1)^2 - 1 = 4k_1^2 + 4k_1 = 4(k_1^2 + k_1)$ (line 10).

The prose leaves unsaid the disjunction (E1) and its reason for covering every case (as in B2, the proof is slightly short of level 3 there), the two $\vee$I steps (silent: each case's last words, "so $4 \mid n^2$" or "so $4 \mid n^2 - 1$", show which disjunct was reached), the $\exists$E closings, the $\vee$E closing and the $\forall$I closing (silent closings). It writes every opening ("let", "if", "say"), both witnesses (k^2 and $k^2 + k$) and both computations.

8.4 Reading a Proof: Self-Explanation

Section 8.3 showed that a level-3 proof leaves steps unsaid by a fixed convention, and that any doubtful stretch can be expanded and checked. That presupposes a reader who notices the doubtful stretch. This section is about becoming that reader. It teaches one habit, **self-explanation**: for each line of a proof, say what makes it true. The habit is taught here once, on one short proof, and the rest of the book relies on it. When a later proof ends its Idea with "*Before reading on: ...*" or "*Self-explain lines 3–5*", this is the passage it means.

8.4.1 Two readers of one proof

Two people read the same proof, line by line, to the end. Both say they understood it. Asked afterwards whether it was correct, one can answer and point to the reason each line holds; the other can only repeat the proof in other words. What separates them is not intelligence or background. It is what each did at every line.

The first asked, at each line, *what makes this true?* and answered by pointing: to earlier lines, to a definition, to a named result, and to the rule that combines them. The second asked *what does this say?* and answered by restating it. Only the first answer can come out wrong, and that is exactly why it is useful. A restatement always succeeds, so it tests nothing.

To self-explain a line is to name its warrant: the earlier lines, definitions and results it comes from, and the rule that gets it from them.

The answer has the shape of a justification column in a derivation. In the derivations of Chapters 1–7 that column was printed beside every line. In level-3 prose it is gone, and self-explanation is the reader writing it back in, at level 1 or 2 of Definition 8.3.1, one line at a time.

8.4.2 The training passage

Here is a short proof. Read it once, then read the two explanations that follow.

Example. Explaining a proof, well and badly

Claim. For every real number $x > 0$, $x + 1/x \geq 2$.

Proof. Let $x > 0$.

1. $(x-1)^2 \geq 0$.
2. So $x^2 - 2x + 1 \geq 0$.
3. Dividing by x, $x - 2 + 1/x \geq 0$.
4. Hence $x + 1/x \geq 2$.

Self-explain each line of the proof.

Solution. **Explained well.**

- *The opening.* "Let $x > 0$" opens a box for an arbitrary x with the assumption $x > 0$: the claim is $\forall x > 0\ (x + 1/x \geq 2)$, and this is its bounded $\forall$I. Everything below may use $x > 0$; so far nothing has.
- *Line 1* is an instance of the school fact that the square of every real number is at least 0, at the term $x - 1$ (a $\forall$E). It uses no earlier line, and in particular not $x > 0$: it holds for every real x.
- *Line 2* follows from line 1 because $(x-1)^2$ and $x^2 - 2x + 1$ are the same number, by expanding the square; one is replaced by the other (=E).
- *Line 3* follows from line 2 by dividing both sides by x. Two things need $x > 0$ here. Dividing by x needs $x \neq 0$, and dividing both sides of an inequality by a number is guaranteed to keep its direction only when that number is positive. Both come from the assumption $x > 0$, and this is the **only** place in the proof where the assumption is used. The left side becomes $(x^2 - 2x + 1)/x = x - 2 + 1/x$, and the right side $0/x = 0$.
- *Line 4* follows from line 3 by adding 2 to both sides, which keeps any inequality. It is the claim for this x; since x was arbitrary, the claim holds for every $x > 0$ (the closing $\forall$I, silent in the proof as written).

Explained badly.

- *Line 1:* we start with a square, which is positive.
- *Line 2:* multiply it out.
- *Line 3:* now we divide by x and it still works.

- *Line 4:* rearranging gives the answer.

Both explanations follow the order of the proof, and the second is not even wrong everywhere; it is shorter, and every sentence after the first describes what happened (the first we return to below). But compare what each one did with line 3. The first says *why* the inequality survives the division, and in saying it, finds the one place where the hypothesis $x > 0$ earns its keep. The second says "it still works", which is a description of the outcome, not a reason.

The difference shows as soon as the proof is wrong. Change the claim to "for every real $x \neq 0$, $x+1/x \geq 2$", and keep all four lines. The bad explanation reads the new proof exactly as happily as the old one: we start with a square, multiply out, divide by x and it still works, rearrange. The good explanation stops at line 3. It needs "the inequality keeps its direction because $x > 0$", and the hypothesis now says only $x \neq 0$; for negative x, division by x reverses the inequality, and line 3 does not follow. The claim is in fact false: at $x = -1$, $x + 1/x = -2 < 2$. A reader who explains warrants finds the gap; a reader who paraphrases cannot, because a paraphrase of a wrong step reads exactly like a paraphrase of a right one.

The bad explanation also slipped in its first line. A square is not always positive, only at least 0: $(x-1)^2 = 0$ at $x = 1$, where the claim holds with equality. Restating invites such slips, because nothing is being checked. Explaining a warrant forces the exact fact, "every square is at least 0", because the explanation has to cite it.

So the first kind of reading does three things the second does not. It **checks** each step, since a warrant can fail to exist. It **locates the hypotheses**, since every use of a hypothesis is a warrant. And it **prepares the reader to reuse the proof**, since the warrants, not the restatements, are what transfer to the next problem.

8.4.3 How to explain a line

For each line, ask four questions, in this order.

1. **What kind of line is it?** An assumption that opens a box ("let", "suppose", "choose … such that"), a line derived from earlier ones, or the closing of a box ("hence", "so", "since x was arbitrary"). In a derivation the annotation answers it: "asm, for $\to$I" beside an assumption names the rule that will close its box (Section 8.1).
2. **Where does it come from?** Name the earlier lines, the definitions and the results it uses. "Line 2", "the definition of even", "Theorem 2.1.3", "the hypothesis $x > 0$". A result is named, not paraphrased.
3. **By which step?** A rule of Chapters 1–7, or a piece of arithmetic that you can say in one clause ("divide both sides by x", "expand the square").
4. **Is it legitimate here?** Is every cited line still available, inside the boxes that are open (Definition 1.4.2)? Is a chosen name fresh? Does a division have a non-zero divisor, and does an inequality keep its direction?

You do not need to write out a derivation; a phrase per line is enough, at the level of the "explained well" column above. When you cannot answer one of the four questions, you have found what to look at: either a step you do not yet understand, or a gap in the proof. Section 8.3 says what to do next: expand that stretch to level 1.

Quick check. Which of the four questions does the bad explanation of line 3 in Example (Explaining a proof, well and badly) fail to answer?

Solution. The fourth, whether the step is legitimate here: does dividing by x keep the direction of the inequality? With it the second fails too, since the warrant $x > 0$ is never named.

Warning. **Self-explanation is not re-reading, and not summarizing.** Reading a proof twice more makes it familiar, not checked: familiarity is what makes the bad explanation of Example (Explaining a proof, well and badly) feel adequate. A summary ("the proof expands a square and divides by x") is useful afterwards, but it skips exactly the warrants that matter, such as why the division by x is allowed and keeps the inequality's direction. Self-explanation is local and backward-looking: each line, against what came before it.

Quick check. Here is a six-line proof that for real numbers x and y, if $x + y > 2$, then $x > 1$ or $y > 1$.

1. We prove the contrapositive: if it is not the case that $x > 1$ or $y > 1$, then $x + y > 2$ fails.
2. Suppose it is not the case that $x > 1$ or $y > 1$.
3. Then $x \leq 1$ and $y \leq 1$.
4. Adding these, $x + y \leq 2$.
5. So $x + y > 2$ fails.
6. This proves the contrapositive, and hence the claim.

Self-explain lines 2–4.

Solution. Line 2 is an assumption: it opens the $\to$I box for the contrapositive announced in line 1, whose hypothesis is $\neg(x > 1 \vee y > 1)$.

Line 3 comes from line 2 in two steps. First Proposition 2.5.5 (a), $\neg(P \vee Q) \vdash \neg P \wedge \neg Q$, gives "not $x > 1$, and not $y > 1$". Then the order fact that $\neg(u > 1)$ means $u \leq 1$, for every real u, turns each conjunct into the form written.

Line 4 uses both conjuncts of line 3 ($\wedge$E, twice) and the order fact that inequalities may be added: $x \leq 1$ and $y \leq 1$ give $x + y \leq 1 + y \leq 1 + 1 = 2$.

(For the record: line 5 is the order fact $x + y \leq 2$ means $\neg(x + y > 2)$; line 6 closes the box by $\to$I and then passes from the contrapositive to the claim by Proposition 2.5.2, a classical step.)

8.4.4 Naming the unstated fact

Section 3.3 turned a paragraph of argument into a sequent: its conclusion, its stated premises, and the unstated facts that close its gaps. A proof is read the same way, one line at a time. Often the second question, *where does it come from?*, has an answer the writer did not write: a fact used but not stated. Self-explaining the line means naming that fact, and the fact then turns out to be one of three things. It may be a school fact or a definition: name it. It may be an earlier result: cite it. Or it may be false, and then the line fails. Two lines that look alike:

- "Since n^2 is a multiple of 3, so is n." The unstated fact, "if $3 \mid n^2$, then $3 \mid n$", is true. A careful proof proves it, by cases on the remainder of n on division by 3.
- "Since n^2 is a multiple of 4, so is n." The unstated fact, "if $4 \mid n^2$, then $4 \mid n$", is false at $n = 2$. Chapter 1 met it as a missing premise (Exercise (B5: Supply the missing premise)).

A reader who only paraphrases reads both lines alike.

Quick check. In "Since $x > 1$, $x^2 > x$", name the unstated fact and the hypothesis it needs.

Solution. Multiplying both sides of $x > 1$ by x keeps the direction of the inequality because $x > 0$; and $x > 0$ follows from $x > 1 > 0$.

8.4.5 Strategies of good readers

Self-explanation is the habit for single lines. Readers who are good at proofs also do a few things with a proof as a whole. Here are five, shown on one short proof.

Claim. For real numbers $x > 0$ and $y > 0$, $\sqrt{xy} \le (x + y)/2$.

Proof. Let $x > 0$ and $y > 0$. Since $x = (\sqrt{x})^2$, $y = (\sqrt{y})^2$ and $\sqrt{xy} = \sqrt{x}\,\sqrt{y}$,

$$x + y - 2\sqrt{xy} = (\sqrt{x} - \sqrt{y})^2 \ge 0.$$

Hence $2\sqrt{xy} \le x + y$, and dividing by 2 gives the claim.

1. **Try it first.** Before reading the proof, spend a few minutes on the claim yourself. A natural attempt squares both sides: $xy \le (x + y)^2/4$ is $4xy \le x^2 + 2xy + y^2$, that is, $0 \le (x - y)^2$, which is true. Already you know what the difficulty is (the square root) and what your attempt rests on (that a square is at least 0).
2. **Identify the framework.** Find the skeleton, level 1 of Definition 8.3.1. Here: a bounded $\forall$I on x and y (opened by "Let", closed silently) with the bounds $x > 0$, $y > 0$, then a direct chain. There is no case split, no contradiction, and no chosen witness. Knowing that tells you where to look: every line is a computation, so every warrant is an algebraic fact.
3. **Break it into sub-proofs.** The proof has one claim inside it, the displayed identity, and then a short finish. A claim inside a proof is proved once and then used, as with a lemma in Section 8.2 ("It suffices to show"). Read the claim on its own: it needs $\sqrt{xy} = \sqrt{x}\sqrt{y}$ (true for $x, y \ge 0$) and $(\sqrt{x})^2 = x$.
4. **Illustrate a hard line with an example.** If the identity is not clear, try numbers. With $x = 1$ and $y = 4$: the left side is $1 + 4 - 2 \cdot 2 = 1$, the right side $(1 - 2)^2 = 1$. An example checks nothing in general, but it shows what the line claims.
5. **Compare with your own attempt.** Your squaring route proves the claim too, and it hides a warrant the given proof avoids: passing back from $xy \le (x + y)^2/4$ to $\sqrt{xy} \le (x + y)/2$ uses the converse of "squaring preserves order": from $a^2 \le b^2$ conclude $a \le b$. This holds when $b \ge 0$, because then $|a| \le b$; here $b = (x + y)/2 > 0$. Without that hypothesis it fails: $a = 1$, $b = -2$. Comparing the two routes shows why the given proof introduces $\sqrt{x}$ and $\sqrt{y}$: it never has to undo a squaring. The comparison also links this claim to Example (Explaining a proof, well and badly): put $y = 1/x$, and $\sqrt{xy} = 1$, so the claim becomes $x + 1/x \ge 2$.

The five strategies work at the level of the whole proof and self-explanation at the level of the

line, and they support each other. Trying first tells you which line will be the hard one; the framework tells you which boxes are open at each line, which is what the fourth question of self-explanation needs.

8.4.6 Exercises

8.4.6.1 A. Check your understanding

Exercise. A1

Here is a proof that if $a > 0$, $b > 0$ and $a^2 = b^2$, then $a = b$.

(1) $a^2 - b^2 = 0$. (2) $(a-b)(a+b) = 0$. (3) $a + b > 0$. (4) Hence $a - b = 0$, that is, $a = b$.

Classify each explanation as one that **justifies** its line or one that **paraphrases** it.

(a) "Line 1 follows from the hypothesis $a^2 = b^2$ by subtracting b^2 from both sides."

(b) "Line 2: then the factored form is zero as well."

(c) "Line 2 is line 1 rewritten with the identity $u^2 - v^2 = (u-v)(u+v)$, at $u = a$ and $v = b$."

(d) "Line 3: now we know that $a + b$ is positive."

(e) "Line 3 holds because $a > 0$ and $b > 0$ are hypotheses, and a sum of two positive numbers is positive."

(f) "Line 4: so $a - b$ is zero, which is what we wanted."

Solution.

(a) Justifies: it names the source (the hypothesis) and the step (subtracting b^2).

(b) Paraphrases: "is zero as well" restates line 2 without saying why the product equals the left side of line 1.

(c) Justifies: it names the identity and the instance.

(d) Paraphrases: it restates line 3 and gives no reason.

(e) Justifies: it names both hypotheses and the fact that combines them.

(f) Paraphrases: it gives no warrant. A justification would say: by line 2 a product is zero, so one of its factors is zero; line 3 says the factor $a + b$ is not zero, being positive; so $a - b = 0$ (Proposition 2.3.4, applied to "$a + b = 0$ or $a - b = 0$" and $a + b \neq 0$).

Exercise. A2: Drill

Answer each part in a phrase. (a) and (b) ask whether an explanation justifies or paraphrases; (c) and (d) which of the four questions is at work; (e) and (f) what kind of line it is; (g)–(i) name the unstated fact; (k) locates a hypothesis; (j) and (l) are about the habits of good readers.

(a) A line: "Since n is even, n^2 is even." An explanation: "With $n = 2k$, $n^2 = 2(2k^2)$, and $2k^2$ is an integer." Does it justify or paraphrase?

(b) A line: "$x^2 - 1 = (x-1)(x+1)$." An explanation: "We factor it." Justify or paraphrase?

(c) Which of the four questions asks "May we divide by $x - 1$ here?"

(d) Which question does the answer "this line opens a box" answer?

(e) What kind of line is "Choose $\delta > 0$ such that ...", and what does its derivation print beside it?

(f) What kind of line is "Since ε was arbitrary, ..."?

(g) Name the unstated fact: "Since $0 < a < b$, $a^2 < b^2$."

(h) Name the unstated fact: "$|x - 2| < 1$, so $x > 1$."

(i) Name the unstated facts: "$x^2 = 9$ and $x > 0$, so $x = 3$."

(j) After three readings, a proof feels clear. Has the reader self-explained it?

(k) "Let $x < 0$. Then $1/x < 0$." Which hypothesis is used, and how?

(l) "Before reading the proof, I tried the claim myself by squaring both sides." Which strategy is this?

Solution.

(a) It justifies: it names the definition of even, the witness and the algebra.

(b) Paraphrase. It says what happened, not which identity makes the line true ($u^2 - v^2 = (u - v)(u + v)$ at $u = x$, $v = 1$).

(c) The fourth: is the step legitimate here (is $x - 1 \neq 0$)?

(d) The first: what kind of line it is.

(e) An assumption that opens an ∃E box with a fresh name; "asm, for ∃E".

(f) The closing of a ∀I box.

(g) Multiplying an inequality by a positive number keeps it: $a \cdot a < a \cdot b$ (multiply $a < b$ by $a > 0$) and $a \cdot b < b \cdot b$ (multiply by $b > 0$).

(h) $|u| < r$ holds exactly when $-r < u < r$; here $-1 < x - 2$.

(i) $x^2 - 9 = (x - 3)(x + 3)$, and a product of real numbers is 0 only if a factor is, so $x = 3$ or $x = -3$; then $x > 0$ rules out -3 (DS).

(j) Not necessarily: familiarity is not a warrant checked line by line (the warning "Self-explanation is not re-reading").

(k) $x < 0$. It gives $x \neq 0$, so $1/x$ exists; and $x \cdot (1/x) = 1 > 0$, so $1/x$ has the sign of x.

(l) Try it first.

8.4.6.2 B. Practice

Exercise. B1: Self-explain a counterexample

Here, with its lines numbered, is the half of Example (The square function on the real line) that shows $f(x) = x^2$ is **not** uniformly continuous on $\mathbb{R}$.

(1) By the negation of uniform continuity, with the bounds $x \in \mathbb{R}$ and $y \in \mathbb{R}$ dropped because they always hold, we must show

$$\exists \varepsilon > 0 \; \forall \delta > 0 \; \exists x \; \exists y \; (|x - y| < \delta \wedge |x^2 - y^2| \geq \varepsilon).$$

(2) Take $\varepsilon = 1$. (3) Let $\delta > 0$ be arbitrary, and take $x := 1/\delta$ and $y := 1/\delta + \delta/2$, which

depend on δ. (4) Then $|x-y| = \delta/2 < \delta$, (5) and, since $y > x > 0$,

$$|x^2 - y^2| = (y-x)(y+x) = \frac{\delta}{2}\Big(\frac{2}{\delta} + \frac{\delta}{2}\Big) = 1 + \frac{\delta^2}{4} \geq 1.$$

(3) Since δ was arbitrary, uniform continuity fails.

Self-explain each line, citing labels where a line uses an earlier result. Say at which lines $\delta > 0$ is used.

Solution. *(1)* The goal "not uniformly continuous" is replaced by the displayed formula: the negation of uniform continuity with no $\neg$ in front of a quantifier (Section 6.2's recipe, Example (Three negations from analysis) (c)). The proof now proves that formula. *(2)* 1 is the witness for "there is $\varepsilon > 0$", and $1 > 0$. It is written first and justified by the end. *(3)* $\delta > 0$ is arbitrary: what follows must work for every δ. The numbers x and y are built from δ, which is allowed because δ is already given; $1/\delta$ exists because $\delta \neq 0$. *(4)* $y - x = \delta/2 > 0$, so $|x-y| = \delta/2$, and $\delta/2 < \delta$ because $\delta > 0$. *(5)* Both numbers are positive and $y > x$, so $|x^2 - y^2| = y^2 - x^2 = (y-x)(y+x)$; multiplying out gives $1 + \delta^2/4$, at least 1 because a square is at least 0. *(6)* δ was arbitrary, so the displayed formula holds with $\varepsilon = 1$, and uniform continuity fails. $\delta > 0$ is used in (3), in (4) twice, and in (5).

Second reading · *The same lines, rule by rule*

(1) A cut: the goal "f is not uniformly continuous" is replaced by the displayed formula, and the proof then proves the formula. The formula is the positive form of the negation of uniform continuity, Example (Three negations from analysis) (c), read with $A = \mathbb{R}$; there the two are shown interderivable, by Proposition 6.2.3, Proposition 6.2.2 and the order fact $\neg(u < \varepsilon)$ iff $u \geq \varepsilon$. The direction used here is from the displayed formula to the negation. The bounds can be dropped because a bounded quantifier $\exists x \in \mathbb{R}$ abbreviates $\exists x\,(x \in \mathbb{R} \wedge \dots)$ (Definition 5.1.1), and over the domain $\mathbb{R}$ the conjunct $x \in \mathbb{R}$ always holds.

(2) The witness for $\exists \varepsilon > 0$: a bounded $\exists$I at the term 1, with its bound $1 > 0$. It is written first, but as a rule it is applied last: the instance it needs, the formula after $\exists \varepsilon > 0$ with 1 for ε, is what lines (3)–(6) prove.

(3) "Let $\delta > 0$ be arbitrary" opens a bounded $\forall$I box flagged δ_0 with the assumption $\delta_0 > 0$. As in Section 5.4, the prose uses the letter δ both for the flag and for the bound variable; below, δ in the arithmetic is the flag δ_0. The terms $1/\delta$ and $1/\delta + \delta/2$ are the witnesses for $\exists x$ and $\exists y$ (two $\exists$I, applied at the end of the box). They may mention δ because δ's box is open; the clause "which depend on δ" records this. The term $1/\delta$ exists because $\delta \neq 0$, from $\delta > 0$.

(4) Arithmetic: $y - x = \delta/2$, which is positive since $\delta > 0$, so $|x-y| = \delta/2$; and $\delta/2 < \delta$ since $\delta > 0$. This is the first conjunct.

(5) Arithmetic. Since $\delta > 0$, $x = 1/\delta > 0$ and $y = x + \delta/2 > x$. Then $x^2 < y^2$, so $|x^2 - y^2| = y^2 - x^2 = (y-x)(y+x)$, with $y - x = \delta/2$ and $y + x = 2/\delta + \delta/2$; multiplying out gives $1 + \delta^2/4$, which is at least 1 because $\delta^2/4 \geq 0$. This is the second conjunct; with (4) it gives the conjunction ($\wedge$I), and the two $\exists$I of (3) then give $\exists x\,\exists y\,(\dots)$.

(6) The closing line of the bounded $\forall$I box flagged δ_0 (Proposition 6.1.1). The side condition holds: the proof has no premises; the only assumption mentioning δ_0 is $\delta_0 > 0$, the assumption

of the box this line closes, and no outer box is open; and δ_0 does not appear in the conclusion $\forall\delta > 0\,\exists x\,\exists y\,(\dots)$, whose δ is a bound variable different from the flag. With (2) this proves the displayed formula, and by the cut (1) uniform continuity fails.

$\delta > 0$ is used in (3) (so that $1/\delta$ exists), in (4) (twice) and in (5) (for $y > x > 0$).

Exercise. B2: The unjustified line

Here is a proof that for real numbers $x < y$, $x^3 < y^3$.

(1) Let $x < y$. (2) Then $y^3 - x^3 = (y - x)(y^2 + xy + x^2)$. (3) Since $x < y$, $y - x > 0$. (4) Also $y^2 + xy + x^2 > 0$. (5) So $y^3 - x^3 > 0$, as a product of two positive numbers. (6) Hence $x^3 < y^3$.

Every line is true. Line 1 opens the box; of the rest, all but one are instances of a routine algebra or order fact, cited or not; one line is not an instance of any fact stated so far, and it needs an argument of its own. Self-explain the proof line by line to find it, and supply that argument as its warrant.

Solution. Line 1 opens the box, with the assumption $x < y$. Line 2 is the identity $(y - x)(y^2 + xy + x^2) = y^3 - x^3$, checked by multiplying out. Line 3 cites line 1 and the order fact $u < v$ iff $v - u > 0$. Line 5 cites lines 2, 3 and 4 and the fact that a product of two positive numbers is positive. Line 6 is line 5 rearranged by the same order fact.

Line 4 is the one without a warrant: it cites nothing, and it is not an instance of any fact stated so far. It is not obvious either, since xy can be negative. A warrant: complete the square,

$$y^2 + xy + x^2 = \left(y + \frac{x}{2}\right)^2 + \frac{3x^2}{4} \geq 0,$$

with equality only if $x = 0$ and $y + x/2 = 0$, that is, only if $x = y = 0$. Line 1 rules that out, since $x < y$ gives $x \neq y$. So $y^2 + xy + x^2 > 0$. The repaired line 4 uses the hypothesis $x < y$ a second time, which the proof as printed hides; a reader who explains only "the bracket is positive" has paraphrased line 4 and missed this.

Exercise. B3: Name the unstated facts

Here is a proof.

Claim. For every real $x \geq 0$, $x^3 + 2 \geq 3x$.

(1) $x^3 - 3x + 2 = (x - 1)^2(x + 2)$. (2) $(x - 1)^2 \geq 0$. (3) $x + 2 > 0$. (4) So $x^3 - 3x + 2 \geq 0$, (5) that is, $x^3 + 2 \geq 3x$.

(a) Name the unstated fact behind each line.

(b) Which line uses the hypothesis $x \geq 0$, and how?

(c) Find the least c such that the same proof, with at most two lines reworded, proves the claim for every $x \geq c$. Show that the claim fails at $x = -3$.

Solution. (a)

- (1) A polynomial identity, checked by expanding: $(x^2 - 2x + 1)(x + 2) = x^3 + 2x^2 - 2x^2 - 4x + x + 2 = x^3 - 3x + 2$.
- (2) Every square is at least 0.
- (3) $x \geq 0$ gives $x + 2 \geq 2 > 0$.
- (4) A product of a non-negative and a positive number is non-negative, applied to (2) and (3), then (1) substituted (=E).
- (5) Adding $3x$ to both sides keeps the inequality.

(b) Line (3), and only there: $x \geq 0$ gives $x + 2 \geq 2 > 0$, which line (4) needs as the positive factor.

(c) $c = -2$. For $x \geq -2$, line (3) becomes "$x + 2 \geq 0$", and line (4) uses "a product of two non-negative numbers is non-negative"; every other line stands. Below -2 the claim is false. At $x = -3$: $(-3)^3 + 2 = -25$, while $3x = -9$, and $-25 \geq -9$ fails. (For every $x < -2$, $(x - 1)^2 > 0$ and $x + 2 < 0$, so the product is negative.)

Marking scheme. (10 points) - 5: (a), one per line, each naming the fact (not restating the line). - 2: (b) line (3), with how $x \geq 0$ enters. - 2: (c) $c = -2$, with the reworded lines (3) and (4). - 1: the failure at $x = -3$ computed.

Fatal errors (0 for the step): explanations that restate the lines; in (c), "the hypothesis can be dropped".

8.4.6.3 C. Going deeper

Exercise. C1: A longer proof

Here is a proof that a sequence has at most one limit, with its lines numbered.

> **Claim.** If $a_n \to L$ and $a_n \to M$, then $L = M$.
>
> (1) Suppose $a_n \to L$ and $a_n \to M$, and suppose, for a contradiction, that $L \neq M$. (2) Let $\varepsilon := |L - M|/2$. Then $\varepsilon > 0$. (3) Choose $N_1 \in \mathbb{N}$ such that $|a_n - L| < \varepsilon$ for every $n \geq N_1$, and $N_2 \in \mathbb{N}$ such that $|a_n - M| < \varepsilon$ for every $n \geq N_2$. (4) Let $n := \max\{N_1, N_2\}$. (5) Then $|L - M| \leq |L - a_n| + |a_n - M| < \varepsilon + \varepsilon = |L - M|$. (6) So $|L - M| < |L - M|$, which is impossible. (7) Hence $L = M$.

Self-explain each line, citing labels where a line uses an earlier result. Say where the assumption $L \neq M$ is used, and which rule closes the box it opens.

Solution (Hint 1). Line (3) applies the definition of a limit at ε. What does that need, and which line supplies it?

Solution (Hint 2). Line (5) uses one named result and two lines of (3). Which?

Solution. *(1)* Two openings: the →I box with the two hypotheses, and inside it a box on $L \neq M$, that is, $\neg(L = M)$; the goal $L = M$ is positive, so this is an RAA box. *(2)* A constant, defined from L and M. It is positive because $L \neq M$ gives $|L - M| > 0$: the one use of the assumption. *(3)* The definition of $a_n \to L$ (as in Example (Three negations from analysis) (a)) applied at ε (∀E at a term that is not the obvious one, with its bound $\varepsilon > 0$ from (2)), then ∃E with the fresh name N_1; the same for M, with a second fresh name N_2 (one name for both would be (N4) of Section 8.6). *(4)* A constant: the larger of two natural numbers, so $n \geq N_1$ and $n \geq N_2$. *(5)* The triangle inequality (Example (The triangle inequality by cases)), with $|L - a_n| = |a_n - L|$; then the two bounds of (3) at this n (∀E at n, →E with $n \geq N_1$ and with $n \geq N_2$); then $\varepsilon + \varepsilon = |L - M|$ by (2). *(6)* The clash: $|L-M| < |L-M|$ against the fact that no number is less than itself (¬E). The line does not mention N_1 or N_2, so the two ∃E boxes close on it. *(7)* RAA: the box assumed $\neg(L = M)$ and the line written after it is $L = M$. The closing of the →I box is silent. The assumption is used once, at (2), and it does work there, so the proof is not in costume; Section 8.2 proved the same claim without opening this box (Example (A sequence has at most one limit)).

Marking scheme. (10 points) - 2: (1) and (7), the two boxes and RAA as the closing rule (not ¬I). - 2: (2), $\varepsilon > 0$ from $L \neq M$, named as the use of the assumption. - 2: (3), the instance named, two fresh names. - 2: (5), the triangle inequality cited and both bounds applied at $n \geq N_1, N_2$. - 2: (6), the clash named, and the ∃E boxes closing because ⊥ does not mention N_1, N_2.

Fatal errors (0 for the step): calling (7) ¬I; applying the definition at ε without its positivity.

Exercise. C2: Challenge: A proof about derivations

Self-explain the proof of Lemma 4.8.3 in Section 4.8: for each of its five numbered steps, and for each claim in the paragraph that follows them, say what it uses and why that is legitimate. Then state the idea of the proof in two sentences.

Solution (Hint 1). Keep Definition 1.4.2 open beside you: the paragraph after the five steps checks, in turn, the premises, the scope rule for the transplanted citations, the rule schemas, the last line of each box, and the availability of what the last line cites.

Solution (Hint 2). Start with step 2. Which result supplies $p \vee \neg p$, what is substituted for its letter, and why may that line cite nothing?

Solution. *Step 1.* The premises of the new list are the members of Δ that D or E uses. Clause (D1) of Definition 1.4.2 places premises first, above every box, and puts no restriction on which statements they are. The conclusion will be $\Delta \vdash \psi$, so only members of Δ are taken. The premises p, $\neg p$ of D and E are therefore left out, unless they also belong to Δ, which does no harm.

Step 2. The line $p \vee \neg p$ is Proposition 2.5.3 with the atom p for the letter P, used as a derived rule (Definition 2.3.1) with no premises, so it cites nothing. This is the one classical step, through LEM.

Step 3. The box opened by p contains the non-premise lines of D with their own boxes; every citation of the premise p is redirected to the assumption, and every citation of a premise in Δ to the premise line of step 1. The two degenerate cases (the last line of D is a premise) are handled separately so that the box still ends in ψ, by repetition (R, clause (D3)) where needed.

Step 4. The same, for E with the assumption $\neg p$.

Step 5. The last line ψ is $\vee$E (Definition 1.6.1), citing the disjunction of step 2 and the two boxes, each opened by one disjunct and ending in ψ.

The verification paragraph. It checks that the list is a derivation. (i) Premises are first and in no box, so they are available everywhere. (ii) Inside each new box, a line of D cites the same lines and boxes as before, renumbered, or a premise line, or the assumption, and each of those is available by the scope rule of Definition 1.4.2: the old boxes have simply been placed inside one new box, which contains both the citing line and the cited one. (iii) Each line is still an instance of its rule's schema, since its statement and the statements it cites are unchanged, as in the proof of Lemma 2.3.3. (iv) The last line of each new box is ψ, as a line of that box itself, which $\vee$E requires. (v) The disjunction and the two boxes are available at the last line, which lies in no box. So the last line is a correct $\vee$E, and $\Delta \vdash \psi$.

The idea in two sentences. The lemma is proof by cases on $p \vee \neg p$: excluded middle supplies the disjunction, and the two given derivations, with their premise p or $\neg p$ turned into the assumption of a box, are the two cases of $\vee$E. All the work is in checking that the transplanted lines still cite only available lines.

Marking scheme. (10 points) - 2: step 2 explained: Proposition 2.5.3 as a derived rule at the atom p, citing nothing, and identified as the classical step. - 2: steps 3 and 4 explained: the redirection of citations of p (or $\neg p$) to the assumption and of members of Δ to the premise lines, and the degenerate cases. - 1: step 5 named as $\vee$E with its three citations. - 3: the verification paragraph explained against Definition 1.4.2: availability by the scope rule, schemas unchanged, each box ending in ψ as its own line. - 2: an idea in two sentences that names both the case split on $p \vee \neg p$ and the transplanting of D and E into the two boxes.

Fatal errors (0 for the step): explanations that restate the steps without naming what they use; calling the case split $\neg$I or RAA.

8.5 Writing a Proof

Section 8.4 trained a reader to say what licenses each line of a proof, as Example (Explaining a proof, well and badly) did line by line. This section turns to the writer. A proof is written for a reader who will check it in exactly that way, so the writer's job is to make each line's warrant visible. The section gives a procedure that leads from a statement to plain prose, the conventions of written mathematics, each a rule of Chapters 1–7 made audible, and four worked examples.

8.5.1 A procedure

A reader checks a proof forwards, one line at a time. A writer finds a proof mostly backwards, from the goal. The two directions are not in conflict: Section 3.1 showed that the search for a derivation is goal-directed, and Definition 1.4.2 says that the derivation found is a list in which each line cites only lines above it. The procedure below keeps both, the search in scratch work

and the derivation on the page. Its steps 3 and 6 use the levels of Definition 8.3.1 and the convention of Definition 8.3.2 from Section 8.3.

Algorithm. **Writing a proof of a statement.**

1. **Write the goal in symbols.** Translate the statement, its hypotheses and its conclusion into a formula, as in Section 6.1, with every quantifier and its bound. Read a letter that is not introduced as universally quantified (Section 5.1). Set up the Givens/Goal table: the hypotheses on the left, the goal on the right.
2. **Unwind the definitions.** Replace each defined word ("bounded", "converges", "irrational", "$\subseteq$") by its definition, in the hypotheses and in the goal. This is $\leftrightarrow$E on a definition.
3. **Read the skeleton off the goal** (Section 3.1, extended in Section 6.3), and write it at level 1: the boxes, their first lines and their last lines. Open the $\exists$E boxes of existential hypotheses early.
4. **Fill the boxes:** backwards in scratch, forwards on the page. The search runs from the goal; the written proof runs from the hypotheses, each line citing lines above it. What the search found goes into the Idea, not into the Proof.
5. **Choose constants last in the search, and write them first on the page,** each with what it depends on.
6. **Write level 3,** leaving silent exactly the steps that Definition 8.3.2 allows, and no others.
7. **Reread as a checker:** for each line, is it true, and by which rule does it follow from which available lines? (Section 8.6 makes this a two-column check.)

Steps 1–3 involve no mathematics, only logical form, and they fix the first and last sentence of every box before anything is invented ("skeleton first"). A picture after the next example shows steps 4 and 5.

Quick check. Which step of the procedure produces the sentence "Let $\varepsilon > 0$" at the start of a proof that $a_n \to L$, and which step keeps it at level 3?

Solution. Step 3. Unwound in step 2, the goal begins $\forall \varepsilon > 0$, so its skeleton is a bounded $\forall$I box whose assumption is $\varepsilon_0 > 0$; "Let $\varepsilon > 0$" is the first line of that box. Step 6 keeps it, because the opening of a $\forall$I box is never silent (Definition 8.3.2). The closing "since ε was arbitrary" may be silent.

8.5.2 The procedure, run once

This example runs all seven steps on a statement not seen before. Sets of reals are used here as Chapter 5 used them, with "$a \in A$" a one-place predicate and $A \subseteq B$ read as "every element of A is an element of B".

Example. A subset of a bounded set

Let A and B be sets of real numbers with $A \subseteq B$. Prove that if B is bounded above, then A is bounded above.

Solution. Steps 1 and 2: the goal in symbols, unwound. Over the domain $\mathbb{R}$, "B is bounded above" is $\exists M\ \forall b \in B\ (b \leq M)$ (Section 6.1), and $A \subseteq B$ is $\forall x \in A\ (x \in B)$. The statement is an implication, so the derivation below is the inside of a $\to$I box: its assumption is the first formula (line 2), the hypothesis is the second, and its goal is $\exists M\ \forall a \in A\ (a \leq M)$.

Givens	Goal
$A \subseteq B$; B bounded above	A bounded above
$\forall x \in A\ (x \in B)$; $\exists M\ \forall b \in B\ (b \leq M)$	$\exists M\ \forall a \in A\ (a \leq M)$
$\forall x \in A\ (x \in B)$; $\forall b \in B\ (b \leq M_0)$ ($\exists$E, M_0 fresh)	$\exists M\ \forall a \in A\ (a \leq M)$
the same	$\forall a \in A\ (a \leq M_0)$ (try the witness M_0)
the same, and $a_0 \in A$ ($\forall$I, a_0 fresh)	$a_0 \leq M_0$
the same, and $a_0 \in B$ ($\forall$E at a_0)	$a_0 \leq M_0$ ✓ ($\forall$E at a_0)

The second row is step 2; each later row is one move of steps 3 and 4; the table's last row is where the search ends.

Step 3: the skeleton. The goal is an existential, so it is postponed: no witness is in sight yet. The assumption "B is bounded above" (line 2) is an existential, so its ∃E box opens at once, with a fresh name M_0 and the assumption $\forall b \in B\ (b \leq M_0)$; the box's last line is the goal.

Step 4: filling the box. Backwards: the goal needs a witness M and a proof of $\forall a \in A\ (a \leq M)$. The only bound in sight is M_0, so try it. Then the new goal is a bounded universal: a box "let $a_0 \in A$" whose last line is $a_0 \leq M_0$. That line comes from the bound on B, instantiated at a_0, which needs $a_0 \in B$; and $a_0 \in B$ comes from $A \subseteq B$ instantiated at a_0. Forwards, the same lines in the order of the list are the derivation below (with the bounded rules of Proposition 6.1.1; each bounded ∀E abbreviates a ∀E and a →E, and the bounded ∀I box a ∀I box with a →I box inside).

1		$\forall x \in A\ x \in B$	
2		$\exists M\ \forall b \in B\ b \leq M$	
3	M_0	$\forall b \in B\ b \leq M_0$	asm, for ∃E
4	a_0	$a_0 \in A$	asm, for ∀I
5	a_0	$a_0 \in B$	∀E 1, 4
6	a_0	$a_0 \leq M_0$	∀E 3, 5
7		$\forall a \in A\ a \leq M_0$	∀I 4–6
8		$\exists M\ \forall a \in A\ a \leq M$	∃I 7
9		$\exists M\ \forall a \in A\ a \leq M$	∃E 2, 3–8

The side conditions hold: a_0 appears in neither line 1 nor line 2, not in line 7 and not in the open assumption on line 3; M_0 appears in none of lines 1, 2 and 9.

Step 5: the constant. The witness for the goal is the bound M_0. The search could choose it only once the ∃E box supplied M_0, and confirm it only at line 6; the page writes it first, in the sentence that chooses it. It depends on B only: its box is opened before the box of a_0, so it cannot depend on a_0, and it must not, because a single bound has to serve every element of A.

Step 6: level 3.

> Suppose B is bounded above, and choose an upper bound M of B (depending on B only). Let $a \in A$. Since $A \subseteq B$, we have $a \in B$, and so $a \leq M$. Hence M is an upper bound of A, and A is bounded above.

The two instances of ∀E (lines 5 and 6) are at the obvious term a, and each appears only as the citation that licenses it ("since $A \subseteq B$", "so", which cites the choice of M). The closing of the ∀I box, "since a was arbitrary", is silent. Two steps that are never silent are on the page: the ∃E with its fresh name ("choose an upper bound M"), and the ∃I with its witness and dependence ("M is an upper bound of A"; "depending on B only"). "Suppose B is bounded above" opens the →I box; its closing is silent.

Step 7: reread. The order of the first two sentences matters. With "let $a \in A$" first and M chosen inside, the ∃I would fall inside the box of a_0, giving only $\forall a \in A\ \exists M\ (a \leq M)$, which

says nothing.

↻ **Second reading** · *What the example used*

The example used $A \subseteq B$ once, at line 5, and "B is bounded above" once, at line 9, by ∃E, whose box assumption is used at line 6. The two warnings below concern steps 4 and 5.

8.5.2.1 Scratch paper and the page

Steps 4 and 5 say that the search runs backwards and the page runs forwards. Here is the example's search beside its page.

Scratch paper (from the goal, backwards)	The page (from the hypotheses, forwards)
1. Goal: an M with $a \leq M$ for every $a \in A$.	1. Suppose B is bounded above, and choose an upper bound M of B.
2. B bounded above gives some bound M: try it.	2. Let $a \in A$.
3. Need $a \leq M$ for an arbitrary $a \in A$.	3. Since $A \subseteq B$, $a \in B$.
4. That needs $a \in B$.	4. So $a \leq M$.
5. $a \in B$ comes from $A \subseteq B$. ✓	5. Hence M is an upper bound of A.

Read the scratch column from the bottom up and it is the page, except for the constant: M is found at scratch line 2, after the goal has been read, but on the page it is chosen in the first sentence, before anything uses it.

↻ **Second reading** · *Why the page runs forwards*

Step 4 as a logical fact. The search of Section 3.1 opens a box for the goal before it knows how the box will be filled, and it postpones an existential goal until a witness appears; that is why it works backwards. But a derivation is a list in which line k cites only lines available at k (Definition 1.4.2), and a written proof is that list in sentences. So the order of the page is forced by the rules, not by taste: "$|f(x) - f(a)| < \varepsilon$ holds if $|x - a| < \varepsilon/2$" points from a later line to an earlier one. The backward reasoning lives in the Idea block.

Step 5 as a logical fact. A chosen constant, such as "let $\delta := \varepsilon/2$", is the witness of an ∃I line. The search finds it last, because only the end of the search shows what it must do. On the page it comes first, right after the names it is built from, because later lines use it. Those names are its dependencies; they must be among the names of the boxes open where it is chosen (Section 6.4).

Warning. **The search is not the proof.** A page that reads "we want $x^2 < 4$; this holds if $x < 2$; so let $x < 2$" has copied the search, and each "holds if" points backwards. A backward step can be false, and the page hides it: "$x < 2$ implies $x^2 < 4$" fails at $x = -3$. In scratch work every backward step must be checked as an implication pointing forwards; on the page

only the forward chain appears, each line citing lines above it.

Warning. **A constant without its dependencies is a claim nobody can check.** In a proof that x^2 is uniformly continuous on $\mathbb{R}$, the sentence "let $\delta := \min\{1, \varepsilon/(2|x|+1)\}$" looks innocent. Asked "what does δ depend on?", the formula answers: on ε **and on** x, so the δ was chosen inside the box of x, and the proof shows only continuity at every point. Uniform continuity is false here (Example (The square function on the real line)). Writing "(depending on ε only)" after the constant is what lets a reader catch the error.

Quick check. On the page, which sentence carries the witness of the final ∃I, and why can the choice of M not come after "Let $a \in A$"?

Solution. "Hence M is an upper bound of A" offers M as the witness; it was chosen in sentence 1. Chosen after "Let $a \in A$", M would sit inside the box of a, and the proof would give only $\forall a \in A\ \exists M\ (a \le M)$, which says nothing (step 7 of the example).

Quick check. A writer's scratch paper reads: "want $|3x-6| < \varepsilon$ when $|x-2| < \delta$; $|3x-6| = 3|x-2|$; so $\delta = \varepsilon/3$." Write the first two sentences of the page.

Solution. "Let $\varepsilon > 0$, and let $\delta := \varepsilon/3$, which is positive and depends on ε only. Let x be a real number with $|x-2| < \delta$."

8.5.3 Conventions of written mathematics

A handful of conventions make a proof checkable. Each is a rule of Chapters 1–7 heard in a sentence.

- **Announce the skeleton.** "We prove the contrapositive", "by induction on n", "we consider two cases", "suppose, for a contradiction, that ...". An announcement tells the reader which box is opening, and so what its last line must be.
- **Name every object when it is introduced,** by "let" (an arbitrary object, ∀I), "choose" (a witness, ∃E, with a fresh name), or "let ... := ..." (a constant, the witness of ∃I), and say what a chosen object depends on.
- **One claim per sentence.** A sentence carries one line, or a short chain of lines joined by "so", each step of which the reader can name. Two unrelated claims joined by "and" hide which one is used later.
- **End with what was shown.** The last sentence of a proof states the goal, or the last line before only closings that Definition 8.3.2 lets be silent. If it is not the statement to be proved, the reader should be told how the two are related.
- **Label displayed claims that will be cited,** (∗) or (1), and cite them by the label. A citation is a line number in prose.

None of these is a matter of style. An object used without being introduced is a line citing nothing, and an announcement that does not match its box (assuming the conclusion of an implication) proves the converse; these are entries (N9) and (N5) of the catalogue in Section 8.6.

8.5.4 Matching sentences and rules

The next two examples go in opposite directions.

Example. Label every sentence

Here is a short proof. Label each sentence with the rule it uses.

> **Claim.** For a real number x, if $|x-1|<1$, then $x>0$.
>
> *Proof.* (1) Assume $|x-1|<1$. (2) Every real number is at least 1 or below it, so either $x \geq 1$ or $x<1$. (3) Suppose first that $x \geq 1$. (4) Then $x \geq 1>0$. (5) Suppose instead that $x<1$. (6) Then $x-1<0$, so $|x-1|=1-x$. (7) Hence $1-x<1$, that is, $x>0$. (8) In either case $x>0$, which proves the claim.

Solution.

Sentence	What it does	Rule
(1)	opens a box with the hypothesis	assumption for $\to$I
(2)	states the disjunction the split needs	a fact about real numbers
(3)	opens the first case	assumption for $\vee$E
(4)		algebra
(5)	opens the second case	assumption for $\vee$E
(6)		algebra
(7)	uses the assumption of (1), whose box is still open	algebra
(8)	combines the cases, then closes the outer box	$\vee$E on (2), (3)–(4), (5)–(7); then $\to$I

Sentence (7) uses the assumption made in (1), and it may: that box is still open.

Measured against Definition 8.3.2, nothing that must be written is missing: the split is on the page with a disjunction that covers every real x, each case opens with its assumption, and what is left out may be silent. Labeling also shows where the work is: the case split exists only because $|\cdot|$ is defined by cases, and the school fact $-|u| \leq u$, used with $u=x-1$, gives $-1<x-1$ in one step, with no $\vee$E at all.

Sentence	What it does	Rule
(1)	opens a box with the hypothesis	assumption for $\to$I
(2)	states a fact about real numbers	a school fact, entered as a line
(3)	opens the first case box	assumption for $\vee$E
(4)	$x \geq 1$ and $1 > 0$ give $x > 0$	$\to$E, with the fact "$a \geq b$ and $b > c$ imply $a > c$"
(5)	opens the second case box	assumption for $\vee$E
(6)	computes $\lvert x - 1 \rvert$	$\to$E twice, with the facts "$x < 1$ implies $x - 1 < 0$" and "$a < 0$ implies $\lvert a \rvert = -a$"
(7)	uses the assumption of (1)	=E, putting (6) into (1), then $\to$E with "$1 - x < 1$ implies $x > 0$"
(8)	combines the cases, then closes the outer box	$\vee$E on (2), (3)–(4), (5)–(7); then $\to$I on (1)–(8)

The facts of school algebra enter as implications, each used by $\to$E. Sentence (8) closes two boxes, which is normal.

Measured against Definition 8.3.2, the proof leaves out nothing that must be written: the split in (2) is on the page, and the disjunction it splits is a school fact (every real number is at least 1 or below it), so it covers every case; each case box opens with its assumption; and every step it leaves unwritten is of a kind that may be silent, for instance the $\forall$E instances of the school facts at the obvious terms, the conclusion $x > 0$ of case 1, which is left to the chain $x \geq 1 > 0$ of (4), the $\wedge$I behind (4), whose conjuncts are displayed, the =E of (7), and the closing of the $\to$I box, for which "which proves the claim" stands. Labeling also shows where the work is. The case split exists only because $\lvert\cdot\rvert$ is defined by cases (Chapter 1); the school fact $-\lvert u \rvert \leq u$, used with $u = x - 1$, gives $-1 < x - 1$ in one step, and a proof written that way would have no $\vee$E at all.

The second direction starts from a derivation whose premises are facts about integers.

Example. From a derivation to a paragraph

Let n be an integer. Write E for "n is even", O for "n is odd", F for "$n + 1$ is even", and G for "$n(n + 1)$ is even". The derivation below uses four facts about integers as premises. Write the paragraph a mathematician would write for it.

1	$E \vee O$	(every integer is even or odd)
2	$E \to G$	(an even factor makes a product even)
3	$O \to F$	(odd plus one is even)
4	$F \to G$	(an even factor again)
5	$\quad E$	asm, for $\vee$E
6	$\quad G$	$\to$E 2, 5
7	$\quad O$	asm, for $\vee$E
8	$\quad F$	$\to$E 3, 7
9	$\quad G$	$\to$E 4, 8
10	G	$\vee$E 1, 5–6, 7–9

Solution.

> Every integer is even or odd, so we consider two cases. If n is even, then $n(n+1)$ has an even factor, and so it is even. If n is odd, then $n+1$ is even, since an odd number plus one is even; so again $n(n+1)$ has an even factor and is even. In either case $n(n+1)$ is even.

Compare the two. The paragraph keeps one sentence for each box and each rule that matters: line 1 becomes "we consider two cases", with its reason; the two boxes become the two "if" sentences; line 10 becomes "in either case". It drops the line numbers and the rule names, and it merges the two $\to$E steps of lines 8 and 9 into one sentence joined by "so". It never skips a box: the reader still sees both cases and sees them combined.

By Definition 8.3.2, the closing "in either case ..." could be dropped at level 3; the split of line 1, with the reason it covers every integer, could not. The paragraph leaves the premises unproved. Premises 2–4 are one line of algebra each (for instance, $n = 2k$ gives $n(n+1) = 2(k(n+1))$), and a careful write-up includes them; premise 1, that every integer is even or odd, is the fact (E1) that this volume takes as known.

8.5.5 Generalizing a proof

A proof written about one object sometimes proves much more. To **generalize** it, replace the object by a fresh name, the move of $\forall$I (Section 5.4), and read every line again. A line that uses only what the new name's box assumes stays. A line that uses a property of the particular object must be proved again, or the generalization fails.

Example. One of three, generalized

Here are two proofs that one of 17, 18, 19 is a multiple of 3. (P1) $18 = 3 \cdot 6$. (P2) $17 = 3 \cdot 5 + 2$, so $17 + 1 = 3 \cdot 6$. Which one generalizes to "for every integer n, one of n, $n+1$, $n+2$ is a multiple of 3"? Write the general proof.

Solution. (P1) does not generalize: it uses a fact about 18 alone. (P2) does, once its one particular fact, $17 = 3 \cdot 5 + 2$, is replaced by what every integer has: $n = 3k$, $n = 3k + 1$ or $n = 3k + 2$ for an integer k, a school fact.

> Let n be an integer [assumption for $\forall$I]. Dividing by 3, there is an integer k with $n = 3k$, $n = 3k + 1$ or $n = 3k + 2$; choose such a k [$\exists$E]. If $n = 3k$, n is a multiple of 3. If $n = 3k + 1$, then $n + 2 = 3(k + 1)$. If $n = 3k + 2$, then $n + 1 = 3(k + 1)$. In every case one of the three is a multiple of 3 [$\vee$E]. Since n was arbitrary, this holds for every integer [$\forall$I].

(P2) was the third case, with $k = 5$.

The test: after renaming, can every line still be justified from what the box of n assumes? A line that computes with the particular number fails the test.

Quick check. A proof that $7^2 - 1$ is divisible by 8 reads "$7^2 - 1 = 48 = 8 \cdot 6$". Does it generalize to every odd n? What would?

Solution. No: it computes with 7. A general proof writes $n = 2k + 1$, so $n^2 - 1 = 4k(k + 1)$, and shows that $k(k + 1)$ is even, since one of k, $k + 1$ is even. That is the proof of Exercise (B2: Back from prose to a skeleton, with blanks).

8.5.6 Exercises

8.5.6.1 A. Check your understanding

Exercise. A1

Here is a proof.

> **Claim.** For real numbers x and y, if $x + y > 2$, then $x > 1$ or $y > 1$.
>
> (1) Assume $x + y > 2$. (2) Suppose, for a contradiction, that it is not the case that $x > 1$ or $y > 1$. (3) Then $x > 1$ fails and $y > 1$ fails. (4) So $x \leq 1$ and $y \leq 1$. (5) Adding, $x + y \leq 2$. (6) This contradicts $x + y > 2$. (7) Hence $x > 1$ or $y > 1$, which proves the claim.

(a) Label each sentence with the rule it uses, as in Example (Label every sentence).

(b) Which sentence closes an RAA box, and which uses one of De Morgan's laws (Proposition 2.5.5)? Which law?

(c) Everything inside the box of (2) follows from (2) alone, until the clash in (6). Which face of "by contradiction" (Section 8.1) does the proof wear, and which face is it really? Write the first sentence of the rewritten proof.

(d) Sentence (7) closes two boxes. Which of the two closings may be silent at level 3, by Definition 8.3.2?

Solution. (a)

Sentence	Rule
(1)	assumption, opening the $\to$I box
(2)	assumption $\neg(x > 1 \lor y > 1)$, opening an RAA box
(3)	Proposition 2.5.5 (a), $\neg(P \lor Q) \vdash \neg P \land \neg Q$, applied to (2)
(4)	$\land$E on (3), twice; $\to$E with the order fact "$\neg(a > 1)$ implies $a \leq 1$", twice; then $\land$I
(5)	$\land$E on (4), then $\to$E with "$a \leq 1$ and $b \leq 1$ imply $a + b \leq 2$"
(6)	$\to$E with the order fact "$a \leq 2$ implies $\neg(a > 2)$" on (5), then $\neg$E on (1) and the result
(7)	RAA on the box (2)–(6); then $\to$I on (1)–(7)

(b) Sentence (7) closes the RAA box: the box opened with the negation of the goal $x > 1 \lor y > 1$ and closes with the goal itself. Sentence (3) uses De Morgan's law (a), a direction derived without RAA.

(c) It wears face (iii), genuine RAA, but it is a contrapositive in costume, as in Example (A contrapositive in costume): the hypothesis (1) is used only at the clash, and everything else derives $x + y \leq 2$ from $x \leq 1$ and $y \leq 1$. The rewrite opens: "We prove the contrapositive: if $x \leq 1$ and $y \leq 1$, then $x + y \leq 2$." The rewrite is clearer, but it does not remove the classical step: passing from the contrapositive back to the claim is Proposition 2.5.2, which uses RAA.

(d) The closing of the $\to$I box ("so if $x + y > 2$, then ...") may be silent. The conclusion of the RAA box, "hence $x > 1$ or $y > 1$", may not.

Exercise. A2: Drill

Answer each part in a line. The parts follow the procedure: (a)–(b) symbols and unwinding, (c)–(d) the skeleton, (e) the order on the page, (f), (g), (i) the conventions, (h) the search and the page, (j)–(k) generalizing, (l) the Givens/Goal table.

(a) Write in symbols, with every quantifier and bound: "the square of every odd integer is odd" (domain $\mathbb{Z}$).

(b) Unwind "the sequence (a_n) is bounded" (Section 5.6).

(c) The goal is $\forall \varepsilon > 0\ \exists \delta > 0\ \forall x\ (|x - a| < \delta \to |f(x) - f(a)| < \varepsilon)$. What are the first and the last line of the outer box?

(d) A hypothesis says "some integer k has $n = 2k + 1$". When does its box open?

(e) On the page, which comes first: "let $\delta := \varepsilon/4$" or "let x be a real number with $|x - 1| < \delta$"?

(f) "x is positive and y is negative, and so $xy < 0$ and $x - y > 0$." Which convention does this break?

(g) A proof uses c in its fourth sentence, and c appears nowhere before. Which convention, and which entry of Section 8.6?

(h) Is "this holds if $x < 2$" a sentence of the page?

(i) Claim: $x^2 - 4x + 3 < 0$ for $1 < x < 3$. A proof ends "... so $x^2 - 4x + 3 = (x-1)(x-3)$." What is missing?

(j) A proof about the number 12 uses only "12 is a multiple of 4". For which integers does the same proof prove the same claim?

(k) A proof about "an arbitrary odd n" uses $n > 0$ in one line. What does it prove?

(l) Givens $\forall x \in A\ (x \in B)$ and $\forall x \in B\ (x \in C)$; Goal $\forall x \in A\ (x \in C)$. Write the next row of the Givens/Goal table.

Solution.

(a) $\forall n\,(\exists k\,(n = 2k+1) \to \exists j\,(n^2 = 2j+1))$.

(b) $\exists M \in \mathbb{R}\ \forall n \in \mathbb{N}\ (|a_n| \le M)$.

(c) First: the assumption $\varepsilon_0 > 0$, "Let $\varepsilon > 0$" (assumption for $\forall$I). Last: $\exists \delta > 0\ \forall x\ (|x - a| < \delta \to |f(x) - f(a)| < \varepsilon_0)$, by $\exists$I with the chosen δ.

(d) At once (step 3): an $\exists$E box with a fresh name k.

(e) The constant (step 5). The box of x needs δ; chosen after x, δ could depend on x.

(f) One claim per sentence: two claims joined by "and" hide which one is used later.

(g) Name every object when it is introduced; (N9), a line that cites nothing.

(h) No. It points backwards, from the goal, and belongs to scratch work or the Idea (step 4).

(i) The end: $x - 1 > 0$ and $x - 3 < 0$, so the product is negative. The last sentence must be the claim (end with what was shown).

(j) For every multiple of 4: rename 12 to a fresh n, with the assumption $4 \mid n$.

(k) Only the claim for positive odd n. For all odd n, the box would have to assume no more than oddness, so that line must be proved again without $n > 0$.

(l) Givens: the same, and $a_0 \in A$ ($\forall$I, a_0 fresh). Goal: $a_0 \in C$.

8.5.6.2 B. Practice

Exercise. B1: From skeleton to prose

Each part gives a statement and a level-1 skeleton of its proof. Write the proof at level 3, keeping every step that Definition 8.3.2 says is never silent.

(a) For every integer n, if n is odd, then $3n + 5$ is even. (Domain $\mathbb{Z}$.)

1	n_0 | $\exists k\,(n_0 = 2\cdot k+1)$	asm, for $\to$I
2	k_0 | $n_0 = 2\cdot k_0+1$	asm, for $\exists$E
3	$3\cdot n_0+5 = 2\cdot(3\cdot k_0+4)$	arith 2
4	$\exists k\,(3\cdot n_0+5 = 2\cdot k)$	$\exists$I 3
5	$\exists k\,(3\cdot n_0+5 = 2\cdot k)$	$\exists$E 1, 2–4
6	$\exists k\,(n_0 = 2\cdot k+1) \to \exists k\,(3\cdot n_0+5 = 2\cdot k)$	$\to$I 1–5
7	$\forall n\,(\exists k\,(n = 2\cdot k+1) \to \exists k\,(3\cdot n+5 = 2\cdot k))$	$\forall$I 1–6

(b) No real number x satisfies $x^2+1 = 2x-1$. (Domain $\mathbb{R}$.)

1	$\exists x\,(x^2+1 = 2\cdot x-1)$	asm, for $\neg$I
2	x_0 | $x_0^2+1 = 2\cdot x_0-1$	asm, for $\exists$E
3	$(x_0-1)^2+1 = 0$	arith 2
4	$(x_0-1)^2 \geq 0$	arith
5	$1 \leq 0$	arith 3, 4
6	$\neg(1 \leq 0)$	arith
7	$\bot$	$\neg$E 5, 6
8	$\bot$	$\exists$E 1, 2–7
9	$\neg\exists x\,(x^2+1 = 2\cdot x-1)$	$\neg$I 1–8

(c) For a real number x: $x^3 = x$ if and only if $x = -1$, $x = 0$ or $x = 1$. (Here x is fixed; "or" chains group to the right.)

1	$x^3 = x$	asm, for $\leftrightarrow$I
2	$(x+1)\cdot(x\cdot(x-1)) = 0$	arith 1
3	$x = -1 \vee (x = 0 \vee x = 1)$	product zero, arith 2
4	$x = -1 \vee (x = 0 \vee x = 1)$	asm, for $\leftrightarrow$I
5	$x = -1$	asm, for $\vee$E
6	$x^3 = x$	arith 5
7	$x = 0 \vee x = 1$	asm, for $\vee$E
8	$x = 0$	asm, for $\vee$E
9	$x^3 = x$	arith 8
10	$x = 1$	asm, for $\vee$E
11	$x^3 = x$	arith 10
12	$x^3 = x$	$\vee$E 7, 8–9, 10–11
13	$x^3 = x$	$\vee$E 4, 5–6, 7–12
14	$x^3 = x \leftrightarrow (x = -1 \vee (x = 0 \vee x = 1))$	$\leftrightarrow$I 1–3, 4–13

Hint: line 3 of (c) uses the school fact that a product of real numbers is zero only if one of its factors is zero.

Solution. (a) Let n be an integer, and suppose n is odd. Choose an integer k with $n = 2k+1$. Then

$$3n + 5 = 6k + 3 + 5 = 2(3k + 4),$$

and $3k + 4$ is an integer, so $3n + 5$ is even.

"Choose an integer k" is the ∃E box, with a fresh name; the witness $3k + 4$ of line 4 is on the page, and it depends on k, hence on n. Lines 5, 6 and 7 close boxes and are silent.

(b) Suppose, to the contrary, that some real x satisfies $x^2 + 1 = 2x - 1$, and fix such an x. Rearranging, $x^2 - 2x + 2 = 0$, that is, $(x - 1)^2 + 1 = 0$. But $(x - 1)^2 \geq 0$, since a square is non-negative, so $1 \leq 0$, which is false. Hence no real x satisfies the equation.

"Suppose, to the contrary" opens the ¬I box of line 1 (the goal is a negation, so no classical rule is used), and "fix such an x" the ∃E box. The clash of lines 5–7 is named, and the conclusion of the ¬I box is written; line 8 is silent.

(c) Let x be real.

(⇒) Suppose $x^3 = x$. Then $(x+1)\,x\,(x-1) = x^3 - x = 0$. A product of real numbers is zero only if a factor is zero, so $x + 1 = 0$, $x = 0$ or $x - 1 = 0$; that is, $x = -1$, $x = 0$ or $x = 1$.

(⇐) Suppose $x = -1$, $x = 0$ or $x = 1$. In the three cases, $x^3 = -1 = x$, $x^3 = 0 = x$ and $x^3 = 1 = x$.

This proves the equivalence.

Both directions are written, and the citation in line 3 is visible. The two splits of (⇐) are merged into "in the three cases"; the disjunction split is the hypothesis itself, and the closing lines 12 and 13 are silent.

Exercise. B2: Put the proof in order

Let (a_n) and (b_n) be sequences of real numbers with $a_n \to L$ and $b_n \to M$, where "$a_n \to L$" is the pattern of Section 6.1: for every $\varepsilon > 0$ there is $N \in \mathbb{N}$ such that $|a_n - L| < \varepsilon$ for every $n \geq N$. Eight of the ten sentences below, in a suitable order, form a level-3 proof that $a_n + b_n \to L + M$; two are distractors. Write down an order for the eight, say which other orders are also admissible, and say why each distractor is wrong.

(i) Let $n \in \mathbb{N}$ with $n \geq N$.

(ii) Since $b_n \to M$ and $\varepsilon/2 > 0$, choose $N_2 \in \mathbb{N}$ (depending on ε) such that $|b_n - M| < \varepsilon/2$ for every $n \geq N_2$.

(iii) Let $N \coloneqq \max\{N_1, N_2\}$, where N_1 and N_2 serve every ε at once.

(iv) Then $n \geq N_1$ and $n \geq N_2$.

(v) Let $\varepsilon > 0$.

(vi) Since $n \geq N$ and $\varepsilon > 0$ were arbitrary, $a_n + b_n \to L + M$.

(vii) Choose $N_1 \in \mathbb{N}$ such that $|a_n - L| < \varepsilon$ for every $n \geq N_1$.

(viii) Let $N \coloneqq \max\{N_1, N_2\}$.

(ix) So $|(a_n + b_n) - (L + M)| \le |a_n - L| + |b_n - M| < \varepsilon/2 + \varepsilon/2 = \varepsilon$.

(x) Since $a_n \to L$ and $\varepsilon/2 > 0$, choose $N_1 \in \mathbb{N}$ (depending on ε) such that $|a_n - L| < \varepsilon/2$ for every $n \ge N_1$.

Solution. An admissible order: (v), (x), (ii), (viii), (i), (iv), (ix), (vi).

The only other admissible order swaps (x) and (ii), which use only ε and one hypothesis each. Nothing else moves: (viii) uses N_1 and N_2; (i) uses N; (iv) uses $n \ge N$ and the definition of N; (ix) uses (iv), (x) and (ii) at this n, and the triangle inequality (Example (The triangle inequality by cases)); (vi) closes the boxes of n and ε, with the $\exists$I of the witness N inside "$a_n + b_n \to L + M$". The maximum exists because there are finitely many numbers.

Distractor (iii) claims that N_1 and N_2 serve every ε at once. They were chosen from $\forall\varepsilon\,\exists N$, inside the box of ε, so they depend on ε and serve only the ε they were chosen for; a uniform N would need $\exists N\,\forall\varepsilon$, which the hypothesis does not give. This is the error that Section 8.6 catalogues as (N2). For $a_n = 1/(n+1)$, $L = 0$, no single N_1 serves every $\varepsilon > 0$: at $\varepsilon = 1/(N_1 + 2)$ the term $a_{N_1} = 1/(N_1 + 1)$ is too far from 0.

Distractor (vii) applies the hypothesis at ε instead of $\varepsilon/2$. It is a legal step, but with it in place of (x), the bound in (ix) becomes $|(a_n + b_n) - (L + M)| < \varepsilon + \varepsilon/2$, not $< \varepsilon$, so (ix) would no longer follow. It is an ε of the wrong kind for the goal.

Exercise. B3: The last step

Here is a level-1 skeleton for the statement "for real numbers x and y, if x is rational and y is irrational, then $x + y$ is irrational", with x and y fixed:

1	$x \in \mathbb{Q} \wedge \neg(y \in \mathbb{Q})$	asm, for $\to$I
2	$x \in \mathbb{Q}$	$\wedge$E 1
3	$\neg(y \in \mathbb{Q})$	$\wedge$E 1
4	$\quad x + y \in \mathbb{Q}$	asm, for $\neg$I
5	$\quad (x + y) - x \in \mathbb{Q}$	difference of rationals 4, 2
6	$\quad y \in \mathbb{Q}$	arith 5
7	$\quad \bot$	$\neg$E 6, 3
8	$\neg(x + y \in \mathbb{Q})$	$\neg$I 4–7
9	$x \in \mathbb{Q} \wedge \neg(y \in \mathbb{Q}) \to \neg(x + y \in \mathbb{Q})$	$\to$I 1–8

and its level-3 prose, with the closing sentence of the $\neg$I box left blank:

> Let x be rational and y irrational. Suppose $x+y$ is rational. Then $y = (x+y)-x$ is a difference of two rational numbers, so y is rational. This contradicts the irrationality of y. _____

Name the move, then complete it.

Solution. The move is skeleton first: the blank is line 8, the line written after the ¬I box of lines 4–7 closes, a line the skeleton fixes before any arithmetic. That box assumed "$x + y$ is rational" and reached $\bot$ at line 7, so line 8 concludes its negation. It is the first face of "by contradiction" (Section 8.1), proof of a negative statement, and uses no classical rule: "irrational" means "not rational", so the goal of the outer box, line 8, is a negation, and the inner box is opened on $x + y \in \mathbb{Q}$ with goal $\bot$.

Completion. "Hence $x + y$ is not rational, that is, $x + y$ is irrational."

The conclusion of a ¬I box is never silent (Definition 8.3.2), so this sentence must be written even though the reader can guess it. Line 9, the closing of the →I box, may stay silent. The sentence before the blank, "this contradicts the irrationality of y", is line 7, the clash of ¬E, which names both lines it cites.

Exercise. B4: Which proof generalizes?

Two proofs that 6 divides $7^3 - 7 = 336$:

(P1) "$336 = 6 \cdot 56$."

(P2) "$7^3 - 7 = 6 \cdot 7 \cdot 8$, a product of three consecutive integers; one of them is even and one is a multiple of 3, so 6 divides the product."

(a) Which proof generalizes to "for every integer n, $6 \mid n^3 - n$", and why does the other not?

(b) Write the general proof at level 3, with its fresh name and every fact it uses.

(c) Which fact about 6 does (P2) leave unstated?

Solution. (a) (P2) generalizes; (P1) computes with 7 and says nothing about another n.

(b) Let n be an integer. Then $n^3 - n = (n-1)\,n\,(n+1)$, a product of three consecutive integers. By Example (One of three, generalized), applied to $n-1$, one of $n-1$, n, $n+1$ is a multiple of 3, so $3 \mid n^3 - n$ (a product with a factor divisible by d is divisible by d: if $a = dc$, then $ab = d(cb)$). By (E1), n is even or odd: if n is even, the factor n is even; if odd, $n-1$ is even. So one factor is even, and $2 \mid n^3 - n$. Since 2 and 3 divide it, so does 6 (by Exercise (B3: Six, two and three, used later), the inclusion $\supseteq$: if $m = 2a = 3b$, then $b = 2(a-b)$ and $m = 6(a-b)$).

(c) "If $2 \mid m$ and $3 \mid m$, then $6 \mid m$", the fact of Exercise (B3: Six, two and three, used later).

Marking scheme. (10 points) - 2: (a) (P2), with the reason (P1) fails. - 2: (b) the factorization and the fresh name. - 2: (b) a multiple of 3 among the factors, cited or proved. - 2: (b) an even factor, from (E1). - 1: (b) the step from 2 and 3 to 6, proved or cited. - 1: (c) the fact named. Fatal errors (0 for the step): a proof that checks finitely many n (N1); "$2 \mid m$ and $3 \mid m$, so $6 \mid m$" with neither proof nor citation (N9).

8.5.6.3 C. Going deeper

Exercise. C1: A constant found backwards

Let $f(x) = 2x + 3$ and let a be a fixed real number. Prove that for every $\varepsilon > 0$ there is $\delta > 0$ such that $|f(x) - f(a)| < \varepsilon$ for every real x with $|x - a| < \delta$. Write your answer as an Idea, which finds δ backwards, followed by a Proof, which runs forwards. Say what δ depends on.

Solution (Hint 1). Compute $|f(x) - f(a)|$ in terms of $|x - a|$ before choosing anything.

Solution (Hint 2). The Proof begins: "Let $\varepsilon > 0$, and let $\delta := \ldots$".

Solution. Idea. We want $|f(x) - f(a)| < \varepsilon$. Compute: $f(x) - f(a) = (2x + 3) - (2a + 3) = 2(x - a)$, so $|f(x) - f(a)| = 2|x - a|$. This is below ε exactly when $|x - a| < \varepsilon/2$. So $\delta = \varepsilon/2$ should work, and it does not involve x or a.

Proof. Let $\varepsilon > 0$, and let $\delta := \varepsilon/2$, which is positive and depends on ε only. Let x be a real number with $|x - a| < \delta$. Then

$$|f(x) - f(a)| = |2(x - a)| = 2|x - a| < 2\delta = \varepsilon.$$

Since x was arbitrary, δ works for ε; since $\varepsilon > 0$ was arbitrary, this proves the claim.

The Idea reads the computation from the goal back to $|x - a|$; the Proof places δ right after ε, before any x appears, and justifies it by the final display. Because δ does not depend on a either, the same proof shows that one δ serves every point: f is uniformly continuous on $\mathbb{R}$, in the sense of Section 6.4.

Marking scheme. (10 points) - 2: an Idea that computes $|f(x) - f(a)| = 2|x - a|$ and reads off δ from the goal. - 2: the Proof opens with $\varepsilon > 0$ arbitrary and introduces δ before any x, with $\delta > 0$ checked. - 2: says what δ depends on (ε only, or ε and a if the choice involves a). - 3: for arbitrary x with $|x - a| < \delta$, the forward chain to $|f(x) - f(a)| < \varepsilon$, each step justified. - 1: a closing sentence with the quantifiers in the order ε, δ, x. Any δ with $0 < \delta \leq \varepsilon/2$ earns full marks. Fatal errors (0 for the step): δ depending on x; a Proof that runs backwards ("this holds if ...").

Exercise. C2: Rewrite it properly

A sequence (a_n) is **bounded** if there is $M \in \mathbb{R}$ with $|a_n| \leq M$ for every $n \in \mathbb{N}$ (Section 5.6). The following proof that a sequence (a_n) with $a_n \to L$ is bounded is correct in substance but badly written.

> We want some M with $|a_n| \leq M$. For big n, a_n is near L, so $|a_n|$ is less than $|L| + 1$ if we use 1 for ε. The small n are only finitely many, so take the biggest. So M is the max of all of them and it works.

(a) Name three faults of the text, each by the convention it breaks (this section) or the step that is not written.

(b) Rewrite the proof at level 3.

Solution (Hint 1). Which objects does the text use without introducing them, and what does each depend on?

Solution (Hint 2). The rewrite begins: "Apply the definition of $a_n \to L$ with $\varepsilon = 1$, and choose $N \in \mathbb{N}$ such that ...".

Solution. *(a)* (i) It writes the search: "we want some M" and "if we use 1 for ε" describe what is being looked for, not lines that follow. (ii) It uses objects without introducing them: "big n" hides an N, obtained from the hypothesis at $\varepsilon = 1$ by ∀E and ∃E, which is never named; "the biggest" does not say of which numbers. (iii) It never says what M is or what it depends on: M is built from the finitely many terms before N, so it depends on N. (Also acceptable: the split into small and big n has no stated reason; the last sentence does not say what was shown.)

(b) Apply the definition of $a_n \to L$ with $\varepsilon = 1$, and choose $N \in \mathbb{N}$ such that $|a_n - L| < 1$ for every $n \geq N$. Let

$$M := \max\{|a_0|, |a_1|, \ldots, |a_{N-1}|, |L| + 1\},$$

the largest of finitely many numbers, so it exists; it depends on N (when $N = 0$ the list is just $|L| + 1$). Let $n \in \mathbb{N}$. Either $n < N$ or $n \geq N$, since every natural number is one or the other. If $n < N$, then $|a_n|$ is in the list, so $|a_n| \leq M$. If $n \geq N$, then by the triangle inequality (Example (The triangle inequality by cases))

$$|a_n| = |(a_n - L) + L| \leq |a_n - L| + |L| < 1 + |L| \leq M.$$

In either case $|a_n| \leq M$. Hence (a_n) is bounded.

The ∀E instance $\varepsilon = 1$ is named, because it is not the obvious one; the witness M of the final ∃I is defined before it is used, with its dependence; the split on $n < N$ or $n \geq N$ carries its reason (every natural number is one or the other).

Marking scheme. (10 points) - 3: (a) three genuine faults, each tied to a convention or an unwritten step (1 each). - 2: (b) applies the hypothesis at a named positive ε and chooses N from it. - 2: defines M before using it, with its dependence on N, bounding the early terms and $|L| + \varepsilon$ (a maximum or a sum). - 2: the two cases $n < N$ and $n \geq N$, the second by the triangle inequality (cited) or an equivalent estimate from $|a_n - L| < \varepsilon$, such as $L - \varepsilon < a_n < L + \varepsilon$ and a split on the sign of a_n. - 1: ends with what was shown. Fatal errors (0 for the step): M chosen before N or as a maximum or sum of infinitely many numbers; the early terms omitted.

Exercise. C3: From $\sqrt{2}$ to $\sqrt{3}$, and not to $\sqrt{4}$

The level-3 proof of Theorem 2.2.1 in Example (The irrationality of the square root of two at four levels) uses (L) and Theorem 2.1.3.

(a) Rewrite it as a proof that no rational number has square 3. You may use (L_3): every rational number is p/q with integers p, q, $q \neq 0$, not both multiples of 3. Prove the lemma your proof needs in place of Theorem 2.1.3.

(b) Rewrite it once more for 4, with (L_4): every rational number is p/q with integers p, q, $q \neq 0$, not both multiples of 4, which is true. Which line fails? Why must some line fail?

(c) Say in one sentence what a generalization keeps and what it must prove again.

Solution (Hint 1). The lemma is "if $3 \mid n^2$, then $3 \mid n$". Split on the remainder of n on division by 3.

Solution (Hint 2). Let r be rational, and suppose $r^2 = 3$. By (L_3), write $r = p/q$ with integers p, q, not both multiples of 3.

Solution. (a) Lemma: if $3 \mid n^2$, *then* $3 \mid n$. Write $n = 3k$, $3k+1$ or $3k+2$. In the first case $3 \mid n$. Otherwise $n^2 = 3(3k^2+2k)+1$ or $n^2 = 3(3k^2+4k+1)+1$, and a number $3j+1$ is not a multiple of 3: from $3j+1 = 3t$ we get $3(t-j) = 1$, and no integer is $\frac{1}{3}$. So those cases clash with $3 \mid n^2$, and $3 \mid n$ holds in each by $\bot$E.

The proof. "Let r be rational, and suppose $r^2 = 3$. By (L_3), write $r = p/q$ with integers p, q, not both multiples of 3. Then $p^2 = 3q^2$, so $3 \mid p^2$, and by the lemma $3 \mid p$. Writing $p = 3m$ gives $9m^2 = 3q^2$, so $q^2 = 3m^2$, and by the lemma $3 \mid q$. So p and q are both multiples of 3, a contradiction; hence $r^2 \neq 3$."

(b) The line "$p^2 = 4q^2$, so $4 \mid p^2$, and by the lemma $4 \mid p$" fails: "if $4 \mid n^2$, then $4 \mid n$" is false at $n = 2$. Some line must fail, because the claim is false: $4 = (2/1)^2$, and $p = 2$, $q = 1$ satisfy (L_4). Lines that are all licensed, from true premises, cannot reach a false conclusion (soundness: Theorem 4.7.2, and the statement of Section 7.2 under "From counter-models to non-derivability" for the quantifier steps).

(c) A generalization keeps the skeleton (the boxes, the choices, the clash) and every line that used only properties shared by the new object; it must prove again each line that used a property of the old one, here the lemma.

Marking scheme. (10 points) - 3: (a) the lemma proved by the residues mod 3, with the clash $3j+1 \neq 3t$ justified. - 2: (a) the rest of the proof, with (L_3) and the lemma used twice. - 2: (b) the failing line named, with the counterexample $n = 2$. - 1: (b) why a line must fail ($\sqrt{4} = 2$ is rational). - 2: (c) skeleton kept, particular lines re-proved. Fatal errors (0 for the step): the lemma asserted without proof (N9); in (b), blaming (L_4).

8.6 Checking Proofs, and What Is Not Allowed

A reader working alone must judge proofs, including their own, and is the person least able to do it. The remedy is procedural. This section gives two questions to ask of every line and a catalogue of the errors that recur, each tied to the rule of Chapters 1–7 that it breaks, and then checks four short proofs line by line.

8.6.1 Two questions for every line

A proof is correct when every line is licensed: by Definition 1.4.2, each line is a premise (a hypothesis, a proved result or a true school fact), an assumption, or one rule applied to available lines. So checking a proof means asking two questions of every line, with the answers in two columns.

- **Is it true?** Is the line correct whenever the hypotheses and the assumptions of the open boxes are true?
- **Does it follow, and by which rule?** From which available lines, and by which rule?

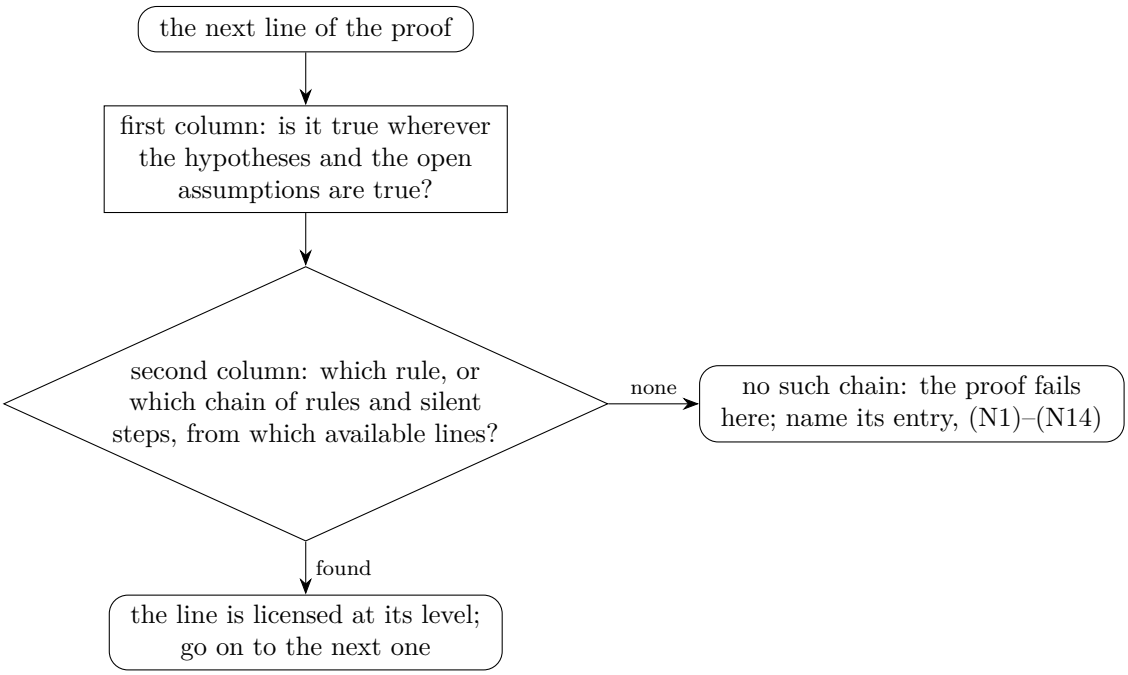
The first column finds where to look; the second decides.

The two questions are different. A line can be **true and not follow**: a correct fact the proof has not earned. This is the dangerous case, because nothing about the line looks wrong.

	It follows	It does not follow
It is true	the normal case: "$n = 2k$, so $n^2 = 2(2k^2)$", in a proof that n^2 is even when n is	the dangerous case: Proof B's last line, "if n^2 is even, then n is even", true and not earned
It is false	the fault is higher up: from the false line "every real x has $x \geq 1$ or $x \leq 0$", $\forall$E at $\frac{1}{2}$ gives "$\frac{1}{2} \geq 1$ or $\frac{1}{2} \leq 0$", which follows and is false	"$x < 1$, so $x^2 < 1$": false at $x = -2$

"False" means false in some situation where the hypotheses and the open assumptions are true. When a line follows and is false, some line it rests on, neither a hypothesis nor an open assumption, is false there too, because the rules never lead from true lines to a false one (Theorem 4.7.2); follow it back until the second column fails.

The two columns, as a procedure for one line:



A "no" in the first column marks a line to look at first; only the second column decides. The silent steps are those of Definition 8.3.2.

↻ **Second reading** · *Two notes on the two columns*

The same holds for derivations with quantifiers and equality, by the soundness statement for predicate logic in Section 7.2 ("From counter-models to non-derivability").

(Proof C's "1 is even" is true given its box's assumption.)

Quick check. A proof that $n^2 \geq n$ for every integer n opens with the line "$n^2 - n = n(n-1) \geq 0$". Is the line true? Does it follow? Would the answers change if n ranged over the reals?

Solution. For an integer n the line is true: no integer lies strictly between 0 and 1, so $n \leq 0$ (both factors ≤ 0) or $n \geq 1$ (both ≥ 0). But as written it follows from nothing: it is the goal rearranged, without the case split that makes it true. For real n it is false at $n = 1/2$, where $n(n-1) = -1/4$. The first column depends on the situation; the second fails in both.

8.6.2 The catalogue

Most wrong proofs fail in one of a few ways, each breaking a particular rule. The entries fall into seven families.

Family	Entries	Slogan
"For all" errors	(N1), (N2), (N3)	A "for all" needs one argument about a name that is new and arbitrary.
Naming errors	(N4), (N12)	A new choice gets a new name, and two names may name one thing.
Implication errors	(N5), (N6)	The box of $P \to Q$ opens with P; Q alone gives no P.
Case errors	(N7), (N8)	Every case, from a disjunction that covers everything.
Scope and citation errors	(N9), (N10)	Cite only what is proved, given, or open around you.
Negation errors	(N11)	Negate exactly: an RAA box opens with the negation of the goal; negate the quantifier, not the bound; $\varphi \wedge \neg\psi$, not an implication.
Iteration and existence errors	(N13), (N14)	An "and so on" needs its step; a "the" needs its object.

Example. What is not allowed

Each entry names the error, the rule it breaks and where that rule lives, and a failing instance whose conclusion is marked **true** or **false**: a wrong argument can reach either.

(N1) **Verifying instances and concluding "for all".** Breaks ∀I, which needs one proof about a fresh name (Definition 5.4.1, Section 5.4). Instance: $n^2 - n + 11$ is prime for $n = 0, 1, \ldots, 10$ (the values are $11, 11, 13, 17, 23, 31, 41, 53, 67, 83, 101$), "so it is prime for every $n \in \mathbb{N}$". **False:** at $n = 11$ the value is $121 = 11^2$.

(N2) **A witness that depends on a variable introduced later, while uniformity is claimed.** Breaks the side conditions of ∃E and ∀I: a witness δ_0 obtained inside the box of x_0 may not leave it (Section 6.4, Example (The search from C to UC)). Instance: "x^2 is continuous at every x, with $\delta = \min\{1, \varepsilon/(2|x| + 1)\}$; so one δ for each ε serves every x, and x^2 is uniformly continuous on $\mathbb{R}$". **False** (Example (The square function on the real line)).

(N3) **Treating the name from ∃E as arbitrary.** Breaks the side condition of ∀I: the witness appears in the open assumption of its box (Definition 5.6.1, Section 5.6). Instance: "Some element of A is positive; let x be such an element. Since x was arbitrary, every element of A is positive." **False** for $A = \{-1, 1\}$.

(N4) **Reusing one witness name for two ∃E's.** Breaks the fresh-name condition: the second name may not appear in an open assumption, and the first box's assumption mentions it (Section 5.6, Example (The sum of two even integers is even)). Sibling boxes may share a name. Instance: "m and n are even, so $m = 2k$ and $n = 2k$ for an integer k; hence $m = n$". **False:** $m = 2$, $n = 4$.

(N5) **Proving the converse.** Breaks →I, whose box opens with the hypothesis, not the conclusion (Definition 1.4.1, Section 1.4). Instance: "If $x^2 > 1$, then $x > 1$: suppose $x > 1$; then $x^2 > x > 1$." **False:** $x = -2$. Argument B of Section 1.1 makes the same move with a **true** conclusion.

(N6) **Affirming the consequent; denying the antecedent.** Breaks Definition 1.4.2 (Section 1.4), being no rule: $P \to Q$ with Q gives no P, and with $\neg P$ gives no $\neg Q$. Each pattern has a refuting row, so no derivation exists (Corollary 4.7.3, Section 4.7): Example (Affirming the consequent is not valid), and for denying the row P false, Q true. Instances: "a multiple of 4 is even and 6 is even, so 6 is a multiple of 4"; "... and 6 is not a multiple of 4, so 6 is not even". Both **false**.

(N7) **A missing case.** Breaks ∨E, which needs one box for each disjunct (Definition 1.6.1, Section 1.6). Instance: "$|x| > 0$ for every real x: every real x has $x > 0$, $x = 0$ or $x < 0$; if $x > 0$, $|x| = x > 0$; if $x < 0$, $|x| = -x > 0$". **False:** the case $x = 0$ was dropped, and $|0| = 0$.

(N8) **Cases that do not exhaust.** Breaks ∨E, whose first input must be a disjunction already derived (Section 1.6). Instance: "$x^2 \geq x$ for every real x: either $x \geq 1$ or $x \leq 0$; in the first case $x^2 = x \cdot x \geq x$, in the second $x^2 \geq 0 \geq x$". **False:** $x = 1/2$ satisfies neither case, and $1/4 < 1/2$.

(N9) **Citing a line that is not available:** the goal itself (circular reasoning), a result proved later, or a general "fact" neither proved nor a true school fact (a false "fact" never qualifies). Breaks Definition 1.4.2: every line is a premise, an assumption, or a rule applied to available lines (Section 1.4). Instances: "if n^2 is even, then n is even: suppose n^2 is even; since n is even, ..." (**true**, and circular); "$x < 1$, so $x^2 < 1$, since squaring preserves $<$" (**false:** $x = -2$).

(N10) **Using an assumption after its box has closed.** Breaks the scope rule of Defini-

tion 1.4.2 (Section 1.4). Instance: "Suppose n is even, say $n = 2k$; then $n^2 = 4k^2$. Hence if n is even, $4 \mid n^2$. Now $n^2 = 4k^2$, so $4 \mid n^2$ for every integer n." **False:** $n = 1$. Proof D below makes the same move with a **true** conclusion.

(N11) **A wrong negation**, in three forms. Each breaks RAA, which closes only a box opened with $\neg\varphi$ (Section 2.5), or the laws of Section 6.2.

- (a) An RAA box opened with a statement that is not the negation of the goal. Instance: "if $x^2 > 9$ and $x > 0$, then $x > 3$: suppose, for a contradiction, that $x < 3$; then $0 < x < 3$ gives $x^2 < 9$, against $x^2 > 9$; hence $x > 3$" (**true**: the box should open with $x \leq 3$, and the case $x = 3$ is never examined).
- (a) A bound negated instead of a quantifier (Proposition 6.2.3). Instance: "$a_n = 0$ does not converge to 0, since some $\varepsilon \leq 0$ has $|a_n - 0| \geq \varepsilon$ for all n" (**false**).
- (a) The negation of $\varphi \to \psi$ written as an implication, $\neg\varphi \to \neg\psi$ or $\varphi \to \neg\psi$, instead of $\varphi \wedge \neg\psi$ (Proposition 6.2.2). Instance: "if $x > 2$, then $x^2 > 4$" is "refuted" because "if $x \leq 2$, then $x^2 \leq 4$" holds at $x = 0$ (**false**: both are true at $x = 0$, so one is not the negation of the other).

(N3) **Distinct letters taken for distinct objects.** Two names introduced separately are treated as naming different objects, though no line says they differ. Breaks $\forall$I: a box opened for two arbitrary names must cover every pair, including the pairs in which both name one object ($\forall$E may put the same term for both, Definition 5.3.1). Instance: "if a and b are roots of $x^2 - 5x + 6 = 0$, then $a + b = 5$: the two roots give $x^2 - 5x + 6 = (x - a)(x - b)$, and comparing coefficients, $a + b = 5$". **False:** $a = b = 2$ are roots, and $2 + 2 = 4$.

(N4) **An "and so on" with nothing behind it.** A step is repeated "and so on", and its outcome is used for every stage, but the passage from each stage to the next, or the reason the process ends, is never proved. Breaks Axiom 4.2.1 (or Theorem 4.2.2): $\to$E on the axiom needs its step as a line, and "and so on" is not that line. Instance: "a non-empty set A of positive rational numbers has a least element: take $a_1 \in A$; if it is not least, take $a_2 \in A$ with $a_2 < a_1$, and so on, until the least element is reached" (**false:** $\{1/n : n \in \mathbb{Z}_{>0}\}$ has no least element, and the process never ends).

(N5) **An object named before it is shown to exist.** A step names an object, with "the ..." or "let ... be the ...", and so presupposes that such an object exists (for "the", exactly one), though no line states or proves it. In (N9) the step asserts a fact that is not available; here the missing premise is never asserted at all, because it hides inside a name. Breaks $\exists$E, which opens its box only by citing a line $\exists x\, \varphi$ (Definition 5.6.1), and the use of "the" (Section 5.8), which needs existence and uniqueness. Instance: "let N be the largest natural number; then $N + 1 \leq N$, so $1 \leq 0$" (**false:** the presupposed premise, that a largest natural number exists, is false).

Second reading · *True conclusions behind an 'and so on' and a 'the'*

Both new kinds of error can reach a true conclusion. The words of (N13)'s instance, said of a non-empty set of natural numbers, reach a true conclusion, but they leave out why the process ends: a strictly decreasing list of natural numbers that starts at a_1 has at most $a_1 + 1$ terms, which needs an induction. "Let m be the least positive integer with $m^2 > 50$; since $7^2 = 49$ and $8^2 = 64$, $m = 8$" reaches a true conclusion too, but the existence of a least such integer is a premise the proof must state: it is (N14). The barber who shaves exactly the villagers who do not shave themselves is (N14) as well; Section 5.6 shows that no such barber can exist

(Theorem 5.6.3).

An entry diagnoses a step, not the conclusion: a **false** instance shows the step leading from true premises to a false conclusion, and a **true** one that a correct conclusion does not excuse it. Some entries name errors met before: (N1) is the first error of Section 5.4's list, and parts (b) and (c) of (N11) are warnings of Section 6.2.

The catalogue lists steps that can lead from true lines to a false one. A step can also be **valid but improper**. "For all reals a and b, $a^2 + b^2 \geq 2ab$; so for all reals b and a, $a^2 + b^2 \geq 2ab$", taken as one line, never leads from truth to falsehood, but no rule of the book writes it in one line, TAUT included (Section 8.3). That is a matter of level (Definition 8.3.1), not an error: at level 0 it is ∀E twice and ∀I twice, and in prose it needs no words.

↻ **Second reading** · *The errors under their traditional names*

Older books name some of these errors as fallacies:

Traditional name	Entry
hasty generalization	(N1)
equivocation: one name for two things	(N4)
affirming the consequent; denying the antecedent	(N6)
false dichotomy	(N8)
begging the question	(N9)
a presupposition ("the . . . ")	(N14)

Affirming a disjunct is a common error of Section 1.6.

Quick check. Which family, and which entry, does this belong to: "Some integer n has $n^2 = n + 6$; fix such an n. Since n was arbitrary, every integer n satisfies $n^2 = n + 6$"?

Solution. A "for all" error, (N3): a name chosen by ∃E is not arbitrary. The premise holds at $n = 3$; the conclusion fails at $n = 0$.

Quick check. Which entry does the sentence "choose N larger than every N_ε, one for each $\varepsilon > 0$" commit, and when is the same move **not** an error?

Solution. (N2). Each N_ε comes from $\forall\varepsilon\,\exists N$, inside the box of its own ε; one N above all of them would be uniform, as if the hypothesis were $\exists N\,\forall\varepsilon$, and usually none exists, since the N_ε can grow without bound as ε shrinks. The move is legitimate when ε ranges over **finitely many** values $\varepsilon_1, \ldots, \varepsilon_k$: then $\max\{N_{\varepsilon_1}, \ldots, N_{\varepsilon_k}\}$ exists, because a finite set of

natural numbers has a maximum.

Warning. **A true conclusion does not make a proof correct.** "I know the result is true" answers the first column for the last line and nothing else. Argument B of Section 1.1 has a true conclusion and only true sentences, and it is wrong (Proof B below).

8.6.3 Four proofs, checked

Here are four short proofs as a student might write them; two are correct and two are not.

Example. Proof A

Claim. For a real number x, if $x^2 - 3x + 2 < 0$, then $1 < x < 2$.

(1) Assume $x^2 - 3x + 2 < 0$. (2) Then $(x-1)(x-2) < 0$. (3) A product of two real numbers is negative only when one factor is positive and the other negative, so either $x - 1 > 0$ and $x - 2 < 0$, or $x - 1 < 0$ and $x - 2 > 0$. (4) In the first case, $1 < x < 2$. (5) In the second case, $x < 1$ and $x > 2$, which is impossible. (6) Hence $1 < x < 2$.

Solution.

Line	True?	Follows? By which rule?
(1)	assumption	opens the $\to$I box
(2)	yes: $(x-1)(x-2) = x^2 - 3x + 2$	=E with the algebraic identity
(3)	yes	$\to$E with the stated fact about signs; a disjunction of two conjunctions
(4)	yes, in this case	$\vee$E, case 1: $\wedge$E gives $x > 1$ and $x < 2$
(5)	the case is impossible	$\vee$E, case 2: $x > 2$ and $x < 1$ give $2 < 1$, which clashes with $\neg(2 < 1)$ ($\neg$E); the case reaches $1 < x < 2$ by $\bot$E
(6)	yes	$\vee$E on (3), (4), (5); then $\to$I on (1)–(6)

Verdict: correct. The disjunction of line (3) is derived from a stated fact, so the cases exhaust and (N8) does not apply. Line (5) is the one to look at twice: "which is impossible" is $\bot$E, which lets a case that cannot occur reach the goal, so that $\vee$E can combine the cases. By Definition 8.3.2 it could be silent, since the clash is stated.

Example. Proof B

Claim. For an integer n, if n^2 is even, then n is even.

(1) Suppose n is even. (2) Then $n = 2k$ for some integer k. (3) So $n^2 = 4k^2 = 2(2k^2)$. (4) Hence n^2 is even. (5) Therefore, if n^2 is even, then n is even.

Solution.

Line	True?	Follows? By which rule?
(1)	assumption	opens a box, with "n is even"
(2)	yes, given (1)	the definition of "even", then $\exists$E with a fresh k
(3)	yes	algebra
(4)	yes, given (1)	$\exists$I with the witness $2k^2$: the definition of "even"
(5)	the statement is true, but ...	**does not follow**: $\to$I on the box (1)–(4) gives "if n is even, then n^2 is even", not (5)

Verdict: wrong; line (5) fails; entry (N5). Every line is true, and lines (1)–(4) prove the converse. $\to$I turns a box from P to Q into $P \to Q$, with the box's assumption on the **left**; writing the claim instead is the step from $P \to Q$ to $Q \to P$. That step has a refuting row, P false and Q true, so no derivation makes it (Corollary 4.7.3); and it fails in practice: "if $x > 1$, then $x^2 > 1$" is true, its converse false at $x = -2$. This is Argument B of Section 1.1, diagnosed in Section 1.4. The claim is true: Theorem 2.1.3 proves it in a box that opens with "n^2 is even".

Example. Proof C

Claim. For an integer n, the number $n^2 - 2$ is not divisible by 4.

(1) Suppose $n^2 - 2$ is divisible by 4, say $n^2 - 2 = 4k$ with k an integer. (2) Then $n^2 = 4k + 2 = 2(2k + 1)$, so n^2 is even. (3) By Theorem 2.1.3, n is even, say $n = 2m$ with m an integer. (4) Then $n^2 = 4m^2$, so $4m^2 = 4k + 2$, and hence $1 = 2(m^2 - k)$. (5) So 1 is even. (6) But 1 is odd, and no integer is both even and odd; this is a contradiction. (7) Hence $n^2 - 2$ is not divisible by 4.

Solution.

Line	True?	Follows? By which rule?
(1)	assumption	opens the $\neg$I box, with "4 divides $n^2 - 2$"; then $\exists$E with a fresh k

(2)	yes, given (1)	algebra, then ∃I with witness $2k+1$
(3)	yes, given (2)	→E with Theorem 2.1.3 and (2); then ∃E with a fresh m
(4)	yes, given (1) and (3)	algebra: subtract $4k$, divide by 2
(5)	yes, given (4)	∃I with witness m^2-k, an integer
(6)	yes	∧I of (5) and "1 is odd", then ¬E against "no integer is both even and odd"
(7)	yes	¬I on the box (1)–(6)

Verdict: correct. The goal is a negation, and the ¬I box opens with the positive statement: a proof "by contradiction" of face (i) of Section 8.1, with no RAA. The witness names k and m differ, as (N4) requires.

Example. Proof D

Claim. For a real number x: if $x > 1$, then $x^2 > x$; and $x^2 + 1 > x$.

(1) Suppose $x > 1$. (2) Multiplying both sides by x, which is positive, gives $x^2 > x$. (3) Hence if $x > 1$, then $x^2 > x$. (4) For the second part, $x^2+1 > x^2$, and $x^2 > x$ by (2), so $x^2 + 1 > x$.

Solution.

Line	True?	Follows? By which rule?
(1)	assumption	opens the →I box for the first conjunct
(2)	yes, given (1)	→E with "multiplying by a positive number keeps >"; $x > 0$ since $x > 1 > 0$
(3)	yes	→I on the box (1)–(2)
(4)	"$x^2 > x$" is **false** for $x = 1/2$	**does not follow**: it cites (2), which lies inside the box closed at (3)

Verdict: wrong; line (4) fails; entry (N10). The claim is a conjunction, so the skeleton is ∧I with two separate proofs. The second cites line (2), derived inside the box that assumed $x > 1$ and closed at line (3), so line (2) is not available. The first column confirms it: at $x = 1/2$, $x^2 = 1/4 < 1/2 = x$, a counterexample to the illegal step, though not to the claim.

The claim is true. For every real x,

$$x^2 - x + 1 = \left(x - \frac{1}{2}\right)^2 + \frac{3}{4} > 0,$$

since a square is non-negative; so $x^2 + 1 > x$, with nothing from the closed box.

In Proofs B and D every sentence sounds right, and in Proof B every sentence is true. The errors are in the connections, which the rules govern: "which rule, from which available lines?" catches both.

Remark (Looking ahead). From *Sets, Relations, and Cardinality*, Chapter 1 on, a warning about one of these errors cites its entry by number, as in "(N3) of the catalogue".

8.6.4 Exercises

8.6.4.1 A. Check your understanding

Exercise. A1

Each line below comes from a proof. Name the catalogue entry of Example (What is not allowed) that it commits, and the rule it breaks.

(a) "We checked $3^n > n^2$ for $n = 0, 1, 2, 3, 4$, so it holds for every $n \in \mathbb{N}$."

(b) "By hypothesis some $x \in A$ has $f(x) = 0$; fix such an x. Since x was arbitrary, $f(x) = 0$ for every $x \in A$."

(c) "To prove that $f(x_1) = f(x_2)$ implies $x_1 = x_2$, suppose $x_1 = x_2$; then $f(x_1) = f(x_2)$."

(d) "To prove $x > 0$, suppose, for a contradiction, that $x < 0$."

(e) "If n is divisible by 6, then n is even. Since n is even, n is divisible by 6."

(f) "Every real x satisfies $x > 0$ or $x < 0$; we consider the two cases."

Solution. (a) (N1): ∀I needs one proof about a fresh name, not five instances.

(b) (N3): x was chosen by ∃E and appears in the open assumption $f(x) = 0$, so ∀I on it breaks the side condition. For $A = \{0, 1\}$ and $f(x) = x$ the conclusion fails.

(c) (N5): the →I box opens with the conclusion, so it proves the converse, which every function satisfies.

(d) (N11) (part (a)): the negation of $x > 0$ is $x \leq 0$, not $x < 0$, so RAA cannot close this box with $x > 0$; the case $x = 0$ is never examined, and ¬I on the box gives only $x \geq 0$.

(e) (N6), affirming the consequent. No rule licenses it, so it breaks Definition 1.4.2: $P \to Q$ with Q gives no P. Here $n = 4$ is even and not divisible by 6.

(f) (N8): the disjunction is not derived, and it is false at $x = 0$, so $\vee$E has no valid first input.

Exercise. A2: Drill

Name the entry of Example (What is not allowed) that each line commits, if any, and say whether its conclusion is true.

(a) "We checked that $n^2 + n$ is even for $n = 1, \dots, 6$, so it is even for every n."

(b) "Let $\varepsilon > 0$ and choose N with $|a_n - L| < \varepsilon$ for $n \geq N$; since ε was arbitrary, this one N works for every ε."

(c) "x and y are odd, so $x = 2k + 1$ and $y = 2k + 1$."

(d) "If n is divisible by 4, then n is even; n is not divisible by 4, so n is not even."

(e) "$n^3 - n$ is even for every integer n: if n is even, say $n = 2k$, then $n^3 - n = 2(4k^3 - k)$; so $n^3 - n$ is even."

(f) "Suppose $x > 0$; then $x^2 > 0$. Hence if $x > 0$, then $x^2 > 0$. Since $x^2 > 0$, ..."

(g) "The negation of 'for every $\varepsilon > 0$, $|x| < \varepsilon$' is 'for some $\varepsilon \leq 0$, $|x| \geq \varepsilon$'."

(h) "To prove 'if n^2 is odd, then n is odd': suppose n is odd; then n^2 is odd."

(i) "Let x and y be solutions of $x^2 = 9$. Then $x^2 - y^2 = 0$, so $(x - y)(x + y) = 0$, and since $x \neq y$, $x + y = 0$."

(j) "$a_1 = 2$ and $a_{n+1} = a_n^2 - 2$. Then $a_2 = 2$, $a_3 = 2$, and so on, so $a_n = 2$ for every $n \geq 1$."

(k) "Let n be the positive integer with $n^2 = 2$. Then $1 < n^2 < 4$, so $1 < n < 2$."

(l) "From '$x > 0$' and '($x > 0$ implies $y > 0$) and ($y > 0$ implies $z > 0$)', conclude '$z > 0$' in one line."

Solution.

(a) (N1); true ($n^2 + n = n(n + 1)$), still not a proof.

(b) (N2); in general false.

(c) (N4); false ($x = 1$, $y = 3$).

(d) (N6), denying the antecedent; false ($n = 2$).

(e) (N7), a missing case: the odd case is never written. The conclusion is true ($n^3 - n = (n - 1)n(n + 1)$, and of $n - 1$, n one is even), still not a proof.

(f) (N10); "$x^2 > 0$" is false at $x = 0$.

(g) (N11), part (b); false. The stated "negation" holds for every x (take $\varepsilon = -1$), also at $x = 0$, where the original holds too, so it is not the negation; the negation is "for some $\varepsilon > 0$, $|x| \geq \varepsilon$".

(h) (N5); the claim is true, but the proof proves the converse.

(i) (N12) (the "since" also cites the unavailable line $x \neq y$, an (N9) step); false: $x = y = 3$ gives $x + y = 6$.

(j) (N13); true. The step "if $a_n = 2$, then $a_{n+1} = 4 - 2 = 2$" is one line, and the proof must write it, as the step of an induction.

(k) (N14). The presupposed premise, that such an integer exists, is false, and so is the conclusion that an integer lies strictly between 1 and 2.

(l) No entry. No primitive rule writes the step in one line: at level 0 it is ∧E twice and →E twice. From Section 6.3 on, a single TAUT line on the propositional skeleton (Proposition 6.3.1), citing both premises, also writes it (Section 8.3). Either way it is a matter of level, not an error.

8.6.4.2 B. Practice

Exercise. B1: Two proofs in two columns

Check each proof in two columns. For a wrong proof, name the first line that fails and its catalogue entry, and say whether the claim is true.

(i) **Claim.** For integers m and n, if m and n are odd, then $m+n$ is even. (1) Assume m and n are odd. (2) Choose integers k and l with $m = 2k+1$ and $n = 2l+1$. (3) Then $m+n = 2k+2l+2 = 2(k+l+1)$. (4) Since $k+l+1$ is an integer, $m+n$ is even.

(ii) **Claim.** For a real number x, if $x^2 < 1$, then $x < 1$. (1) Assume $x^2 < 1$. (2) Suppose, for a contradiction, that $x > 1$. (3) Then $x > 0$, so $x^2 = x \cdot x > x > 1$. (4) This contradicts $x^2 < 1$. (5) Hence $x < 1$.

Solution. (i)

Line	True?	Follows? By which rule?
(1)	assumption	opens the →I box
(2)	yes, given (1)	∧E, the definition of "odd", and two ∃E's with the **different** fresh names k and l
(3)	yes	algebra, =E in a chain
(4)	yes	∃I with the witness $k+l+1$; then the boxes close, and →I

Verdict: correct. In (2) the second witness has a new name, as (N4) of Example (What is not allowed) requires.

(ii)

Line	True?	Follows? By which rule?
(1)	assumption	opens the →I box
(2)	assumption	opens a box with "$x > 1$", announced as RAA
(3)	yes, given (2)	→E with order facts
(4)	yes	(3) gives $\neg(x^2 < 1)$ by order facts; ¬E on (1) and that line
(5)	true, given (1), but ...	**does not follow**: the box assumed $x > 1$, which is not the negation of $x < 1$; ¬I on it gives only $\neg(x > 1)$, that is, $x \leq 1$

Verdict: wrong; line (5) fails, because the box of line (2) opened with the wrong statement; entry (N11). The negation of $x < 1$ is $x \geq 1$, and the case $x = 1$ is never examined: $x = 1$ has $\neg(x > 1)$ and not $x < 1$, a counterexample to the step. The claim is true. Repair: suppose, for a contradiction, that $x \geq 1$; then $x^2 = x \cdot x \geq x \geq 1$, which contradicts $x^2 < 1$; hence $x < 1$.

Exercise. B2: Find the error

Find the error in the following proof, name its catalogue entry, and decide whether the claim is true.

> **Claim.** For all real numbers x and y, $|x| + |y| = |x + y|$.
>
> (1) Let x and y be real numbers. (2) Either x and y are both ≥ 0, or they are both < 0. (3) If both are ≥ 0, then $x + y \geq 0$, so $|x| + |y| = x + y = |x + y|$. (4) If both are < 0, then $x + y < 0$, so $|x| + |y| = -x - y = -(x + y) = |x + y|$. (5) In either case $|x| + |y| = |x + y|$.

Solution. Line (2) fails. Its disjunction is not derived, and it is false when x and y have opposite signs, so line (5) combines cases that do not exhaust: entry (N8) of Example (What is not allowed). Lines (3) and (4) are correct in their cases.

The claim is **false**: for $x = 1$, $y = -1$, $|x| + |y| = 2$ and $|x + y| = 0$; this pair satisfies neither case of line (2). What is true is $|x + y| \leq |x| + |y|$ (Example (The triangle inequality by cases)).

Exercise. B3: Correct, or which entry?

Each proof below is correct, or commits one of (N12), (N13), (N14); you are not told how many of each. For a correct proof, name its most delicate step. For a wrong one, name the entry, say whether the claim is true, and give a counterexample to the claim or a repair of the proof.

(i) *Claim:* for all integers m, n, if m and n are odd, then mn is odd. *Proof:* write $m = 2k+1$ and $n = 2l+1$ with integers k, l; then $mn = 4kl + 2k + 2l + 1 = 2(2kl + k + l) + 1$.

(ii) *Claim:* for all real a, b, if $a^2 = b^2$, then $a = b$ or $a = -b$. *Proof:* $(a - b)(a + b) = a^2 - b^2 = 0$, and a product of real numbers is 0 only if a factor is.

(iii) *Claim:* for all integers m, n, $m^2 + n^2 \geq 1$. *Proof:* two different integers are not both 0, so one of m^2, n^2 is at least 1.

(iv) *Claim:* every integer $n \geq 2$ is a product of primes. *Proof:* if n is prime, it is a product of one prime. If not, $n = ab$ with $1 < a, b < n$; split a and b in the same way, and so on, until only primes remain.

(v) *Claim:* 1 is a solution of $x^2 + x + 1 = 0$. *Proof:* let x be the real solution. Then $x^3 - 1 = (x - 1)(x^2 + x + 1) = 0$, so $x^3 = 1$, and since x is real, $x = 1$.

Solution. (i) **Correct.** The delicate step is the second ∃E, which uses a second name l. The case $m = n$ is covered, since nothing stops $k = l$: the names are distinct, the objects need not be.

(ii) **Correct.** The delicate step is the product fact. It gives "$a - b = 0$ or $a + b = 0$" (→E, citing it), which is the claim, and no distinctness is assumed.

(iii) **(N12); the claim is false**: $m = n = 0$ gives 0. The proof took m and n to be different integers.

(iv) **(N13); the claim is true.** The "and so on" hides why the splitting ends: each factor is at least 2 and smaller than the number split. Repair by strong induction (Theorem 4.2.2): if every k with $2 \leq k < n$ is a product of primes, so is n, being prime or ab with a, b in that range.

(v) **(N14); the claim is false**: $1+1+1 = 3 \neq 0$. The presupposed premise "$x^2+x+1 = 0$ has a real solution" is false, since $x^2 + x + 1 = (x + \frac{1}{2})^2 + \frac{3}{4} > 0$.

Look-alikes: (i) and (iii) both use two letters; only (iii) assumes the letters differ.

Marking scheme. (10 points: 2 per part) - 1: the verdict. - 1: for a correct proof, the delicate step named; for a wrong one, the entry with a counterexample (iii, v) or a repair (iv). Fatal errors (0 for that part): calling (i) an instance of (N12) because k and l might be equal; calling (iv) correct.

8.6.4.3 C. Going deeper

Exercise. C1: The second half

Let $f : \mathbb{R} \to \mathbb{R}$ be continuous at a, as the quantifier pattern of Section 6.1: for every $\varepsilon > 0$ there is $\delta > 0$ such that $|f(x) - f(a)| < \varepsilon$ for every real x with $|x - a| < \delta$. Suppose $f(a) > 0$. Here is a level-1 skeleton for the statement "there is $\delta > 0$ such that $f(x) > 0$ for every x with $|x - a| < \delta$":

1		$\forall \varepsilon > 0\ \exists \delta > 0\ \forall x\ (\lvert x - a\rvert < \delta \to \lvert f(x) - f(a)\rvert < \varepsilon)$	
2		$f(a) > 0$	
3		$f(a)/2 > 0$	arith 2
4		$\exists \delta > 0\ \forall x\ (\lvert x - a\rvert < \delta \to \lvert f(x) - f(a)\rvert < f(a)/2)$	$\forall$E 1, 3
5	δ_0	$\delta_0 > 0$	asm, for $\exists$E
6		$\forall x\ (\lvert x - a\rvert < \delta_0 \to \lvert f(x) - f(a)\rvert < f(a)/2)$	asm, for $\exists$E
7	x_0	$\lvert x_0 - a\rvert < \delta_0$	asm, for $\to$I
8		$\lvert x_0 - a\rvert < \delta_0 \to \lvert f(x_0) - f(a)\rvert < f(a)/2$	$\forall$E 6
9		$\lvert f(x_0) - f(a)\rvert < f(a)/2$	$\to$E 8, 7
10		$f(x_0) > 0$	arith 2, 9
11		$\lvert x_0 - a\rvert < \delta_0 \to f(x_0) > 0$	$\to$I 7–10
12		$\forall x\ (\lvert x - a\rvert < \delta_0 \to f(x) > 0)$	$\forall$I 7–11
13		$\exists \delta > 0\ \forall x\ (\lvert x - a\rvert < \delta \to f(x) > 0)$	$\exists$I 5, 12
14		$\exists \delta > 0\ \forall x\ (\lvert x - a\rvert < \delta \to f(x) > 0)$	$\exists$E 4, 5–13

Its level-3 prose begins:

> Since $f(a) > 0$, the number $f(a)/2$ is positive. By the continuity of f at a, applied with $\varepsilon = f(a)/2$, choose $\delta > 0$ such that $\lvert f(x) - f(a)\rvert < f(a)/2$ for every real x with $\lvert x - a\rvert < \delta$.

The rest is blank. *Name the move, then complete it.* Say which skeleton lines your sentences carry, and what δ depends on.

Solution (Hint 1). What is left is the box of x_0 and the witness.

Solution (Hint 2). "Let x be a real number with $\lvert x - a\rvert < \delta$. Then ..."

Solution. The move: skeleton first: the goal fixes the $\exists$I of line 13 and the box of x_0. Filling it: $\forall$E at the non-obvious instance $\varepsilon = f(a)/2$ (line 4), $\exists$E on its δ (line 5), and that δ as the witness.

Completion.

> Let x be a real number with $\lvert x - a\rvert < \delta$. Then $\lvert f(x) - f(a)\rvert < f(a)/2$, so, as $-\lvert u\rvert \le u$, $f(x) - f(a) > -f(a)/2$, that is, $f(x) > f(a)/2 > 0$. So this δ, which depends only on a, is as required.

"Let x ... with $\lvert x - a\rvert < \delta$" opens the two boxes at line 7; "then $\lvert f(x) - f(a)\rvert < f(a)/2$" is lines 8–9, the $\forall$E at the obvious term x silent; "so ... $f(x) > f(a)/2 > 0$" is line 10, which uses line 2. Lines 11 and 12 are silent closings, and the last sentence is lines 13–14, with the

δ already chosen as witness. It depends on a (through $f(a)$ and the continuity at a) and on no x: the box of δ_0 encloses the box of x_0.

Marking scheme. (10 points) - 2: names the move (skeleton first; a hypothesis applied at a chosen ε, then its δ reused as the witness), or the rules $\forall$E, $\exists$E, $\exists$I. - 3: for an arbitrary x with $|x-a|<\delta$, derives $f(x)>f(a)/2$, with the step from $|u|<r$ to $u>-r$ justified. - 2: concludes $f(x)>0$ using $f(a)>0$, and states the conclusion with δ before x. - 2: matches the sentences to the skeleton's lines. - 1: says δ depends on a and not on x. Second route: refuting $f(x)\le 0$ by $|f(x)-f(a)|\ge f(a)-f(x)\ge f(a)>f(a)/2$ earns the second line and the $f(x)>0$ part of the third. Fatal errors (0 for the step): choosing δ after x; applying continuity at an ε not shown to be positive.

Exercise. C2: A true conclusion

Here is a proof that $\sqrt{2}+\sqrt{3}$ is irrational, a true conclusion. Find the line that fails and its catalogue entry, give a counterexample to the illegal step, and write a correct proof.

> (1) Suppose $\sqrt{2}+\sqrt{3}=r$ for some rational number r. (2) Squaring, $r^2=2+2\sqrt{6}+3=5+2\sqrt{6}$. (3) Hence $\sqrt{6}=(r^2-5)/2$, which is rational, since r is. (4) But $\sqrt{6}=\sqrt{2}\cdot\sqrt{3}$ is irrational, being a product of two irrational numbers. (5) This contradiction shows that $\sqrt{2}+\sqrt{3}$ is irrational.

Solution (Hint 1). Run the second column on every line.

Solution (Hint 2). For a correct proof, isolate $\sqrt{3}=r-\sqrt{2}$ before squaring.

Solution. The failing line is (4). Its first half is true, $\sqrt{6}=\sqrt{2}\cdot\sqrt{3}$, but its reason is not available. "A product of two irrational numbers is irrational" is neither proved nor a school fact: it is false, since $\sqrt{2}$ is irrational (Theorem 2.2.1), and $\sqrt{2}\cdot\sqrt{2}=2$ is rational. So the line cites nothing: entry (N9) of Example (What is not allowed). The product $\sqrt{2}\cdot\sqrt{2}=2$ is a counterexample to the illegal step; there is none to the conclusion, which is true. Lines (1)–(3) and (5), and the $\neg$I skeleton, are correct.

A correct proof. Suppose $\sqrt{2}+\sqrt{3}=r$ with r rational. Then $r>0$, since both square roots are positive. Isolating, $\sqrt{3}=r-\sqrt{2}$, and squaring,

$$3=r^2-2r\sqrt{2}+2, \qquad \text{so} \qquad \sqrt{2}=\frac{r^2-1}{2r},$$

where dividing by $2r$ is legitimate because $r>0$. The right-hand side is rational, which contradicts Theorem 2.2.1. Hence $\sqrt{2}+\sqrt{3}$ is irrational.

Marking scheme. (10 points) - 3: names line (4) as the failing line, and the illegal step as the unproved general "fact", entry (N9). - 2: a counterexample to that step ($\sqrt{2}\cdot\sqrt{2}=2$, or another product of two irrationals that is rational). - 1: says the conclusion is true, so no counterexample to it exists. - 4: a correct proof: the $\neg$I box on "r rational", the algebra,

the division justified by $r \neq 0$, and the clash with Theorem 2.2.1 (or with $\sqrt{3}$ or $\sqrt{6}$, its irrationality proved). Fatal errors (0 for the step): "correcting" line (4) by another unproved fact; dividing by r without $r \neq 0$.

Exercise. C3: Exit check: four student proofs

Judge each proof. If it is wrong, give the line that fails, its entry in Example (What is not allowed), and a counterexample to the illegal step. If it is correct, name the rule of its most delicate step and check that step.

(1) **Claim.** For every integer n, if n is a multiple of 4 or a multiple of 6, then n is even. (1) Let n be an integer that is a multiple of 4 or of 6. (2) Case 1: n is a multiple of 4. Choose an integer k with $n = 4k$; then $n = 2(2k)$, so n is even. (3) Case 2: n is a multiple of 6. Choose an integer k with $n = 6k$; then $n = 2(3k)$, so n is even. (4) In either case, n is even.

(2) **Claim.** For integers m and n, if m and n are even, then mn is divisible by 4. (1) Assume m and n are even. (2) Then $m = 2k$ and $n = 2k$ for an integer k. (3) So $mn = 4k^2$, which is divisible by 4.

(3) **Claim.** For every real x, $|x - 1| \leq |x| + 1$. (1) Either $x \geq 1$ or $x \leq 0$. (2) If $x \geq 1$, then $|x - 1| = x - 1 \leq x + 1 = |x| + 1$. (3) If $x \leq 0$, then $|x - 1| = 1 - x = 1 + |x|$. (4) In either case, $|x - 1| \leq |x| + 1$.

(4) **Claim.** For every real x, if $x^2 = 2x$, then $x = 2$. (1) Assume $x^2 = 2x$. (2) Canceling x from both sides, $x = 2$. (3) Hence if $x^2 = 2x$, then $x = 2$.

Solution (Hint 1). Run the two-column check on every line, and keep the two questions apart: a true claim can have a wrong proof.

Solution (Hint 2). In (1), ask which boxes are open when the second k is chosen. In (3), ask what $\vee$E needs from line (1).

Solution. *(1)* **Correct.** The skeleton is $\forall$I and $\to$I, then $\vee$E on the hypothesis itself, so the cases exhaust. The delicate step is the second "choose an integer k", an $\exists$E reusing a name. It is legal: the first box of k lies inside Case 1, which closed before Case 2 opened, so k appears in no open assumption, not in the existential and not in the conclusion. Sibling boxes may share a name (Section 5.6); (N4) forbids only a name still in use.

(2) **Wrong; line (2) fails; entry (N4) of Example (What is not allowed).** The second witness is named k while the box of $m = 2k$ is open. Counterexample to the step: $m = 2$, $n = 4$; no integer k has $2 = 2k$ and $4 = 2k$. The conclusion is true: with $m = 2k$ and $n = 2l$, $mn = 4kl$.

(3) **Wrong; line (1) fails; entry (N8).** "Either $x \geq 1$ or $x \leq 0$" is not derived, and is false at $x = 1/2$, a counterexample to the step. The conclusion is true; the missing case repairs the proof: if $0 < x < 1$, then $|x - 1| = 1 - x < 1 < |x| + 1$.

(4) **Wrong; line (2) fails; entry (N9); the claim is false.** Canceling uses the "fact" "if

$ax = bx$, then $a = b$", which is neither proved nor a school fact, and is false: $a = 1$, $b = 2$, $x = 0$. It needs $x \neq 0$, which no line provides. The claim fails at $x = 0$; what is true is "$x = 0$ or $x = 2$", from $x(x - 2) = 0$.

Marking scheme. (10 points: 2.5 per proof) - For each of (2), (3), (4): the verdict "wrong" (1); the failing line and its entry, (N4), (N8), (N9) (1); a counterexample to the illegal step (0.5). - For (1): the verdict "correct" (1); the delicate step named as an ∃E reusing a name in a sibling box (or as the ∨E on the hypothesis) (1); that step checked against the side condition (0.5). For (4), "→E without its antecedent $x \neq 0$" is also accepted. Fatal errors (0 for that proof): calling proof (1) wrong as an instance of (N4) because the name k recurs; judging (2) or (3) correct because its conclusion is true.

8.6.4.4 R. Review

Exercise. R1: A dialogue

Two students discuss a claim.

> **Ana:** For every integer n, $n^2 + n$ is even.
>
> **Ben:** Why?
>
> **Ana:** Take $n = 3$: $9 + 3 = 12$. Take $n = 4$: $16 + 4 = 20$. It always works.
>
> **Ben:** That is two numbers.
>
> **Ana:** All right. $n^2 + n = n(n + 1)$. If n is even, the product has an even factor. If n is odd, $n + 1$ is even, so again it has an even factor. Either way $n^2 + n$ is even.
>
> **Ben:** You only did positive numbers.
>
> **Ana:** I never used that n is positive.

(a) Which entry of Example (What is not allowed) does Ana's first argument commit?
(b) Write her second argument as a level-3 proof, and name the rule each sentence carries.
(c) Is Ben's last objection right?
(d) On which fact does the case split rest, and where was it stated?

Solution. *(a)* (N1) of Example (What is not allowed): two instances, then "always".

(b) "Let n be an integer [∀I]. Then $n^2+n = n(n+1)$ [algebra]. By (E1), n is even or odd [∨E, with the reason the cases exhaust]. If n is even, say $n = 2k$ [∃E], then $n(n+1) = 2(k(n+1))$ is even [∃I, witness $k(n + 1)$]. If n is odd, say $n = 2k + 1$, then $n + 1 = 2(k + 1)$, and $n(n + 1) = 2(n(k + 1))$ is even [∃E, ∃I]. In either case $n^2 + n$ is even [∨E, closing; silent at level 3]."

(c) No. The box of n assumed only that n is an integer, and no line used more, so ∀I (Definition 5.4.1, Section 5.4) gives the claim for every integer: n is in no premise, and the conclusion is about the bound variable. A proof that used $n > 0$ would have had to assume it in the box.

(d) (E1), "every integer is even or odd", stated in Section 2.1 as a school fact, before Theo-

rem 2.1.3.

8.7 The Dispute over Excluded Middle

Here is a proof that two numbers with a certain property exist which, when it is finished, still does not tell you what the numbers are. It is correct by the rules of this book. Some mathematicians, called **intuitionists**, reject it, for a century now and for reasons that are mathematical rather than merely fussy. This section shows the proof, says what those mathematicians ask of a proof, finds the one rule that fails to provide it, and asks what can be proved without that rule and what analysis would lose.

8.7.1 Irrational powers, with and without a witness

Here is the standard example of a proof that establishes existence without saying what exists.

Theorem 8.7.1. An Irrational Power of an Irrational Number Can Be Rational

Optional: nothing later depends on this.

There are irrational numbers a and b such that a^b is rational.

Try it first (ten minutes): find such a and b explicitly, and prove that your a and b are irrational and that a^b is rational.

Idea. The first candidates anyone tries are $a = b = \sqrt{2}$. Then the question is whether $\sqrt{2}^{\sqrt{2}}$ is rational, and the attempt stalls there: nothing in this book decides it. There are two ways out. The classical one splits on the undecided question with LEM. If $\sqrt{2}^{\sqrt{2}}$ is rational, the first candidates work. If not, use it as the new a and raise it to $\sqrt{2}$ once more, which turns the exponent into $\sqrt{2} \cdot \sqrt{2} = 2$. The proof below takes this way. The other way chooses an exponent that undoes the square root, and the proposition after the proof takes it.

Proof. Let $s := \sqrt{2}^{\sqrt{2}}$, a positive real number. By excluded middle, s is rational or s is not rational [LEM], and we argue by cases [$\vee$E].

Case 1: s is rational. Take $a := \sqrt{2}$ and $b := \sqrt{2}$. Both are irrational by Theorem 2.2.1, and $a^b = s$ is rational [$\exists$I].

Case 2: s is not rational. Take $a := s$ and $b := \sqrt{2}$. Then a is irrational by the case assumption and b by Theorem 2.2.1, and

$$a^b = (\sqrt{2}^{\sqrt{2}})^{\sqrt{2}} = \sqrt{2}^{\sqrt{2}\cdot\sqrt{2}} = \sqrt{2}^{\,2} = 2,$$

which is rational [$\exists$I].

In either case there are irrational a and b with a^b rational. This proves the theorem.

■

The classical step is the first line, LEM on the statement "s is rational". Every other step uses only rules other than RAA. The proof does **not** tell us which pair works. It hands over two candidate pairs, $(\sqrt{2}, \sqrt{2})$ and $(\sqrt{2}^{\sqrt{2}}, \sqrt{2})$, and a guarantee that one of them succeeds. The next subsection says why some mathematicians do not accept such a proof. Section 2.5 described excluded middle as making a decision; here the decision about "s is rational" is made, and nobody learns its outcome.

Remark. In fact the second case holds: $\sqrt{2}^{\sqrt{2}}$ is transcendental, so in particular irrational, by the Gelfond–Schneider theorem, proved independently by Gelfond and by Schneider in 1934. That theorem lies far beyond this book, and nothing here uses it. It also does not repair the proof above, which never needed to know the answer.

The other way out finds the witness. Since $\sqrt{2}^{\,b} = 2^{b/2}$, the exponent $b = 2\log_2 m = \log_2(m^2)$ gives $\sqrt{2}^{\,b} = m$ for every integer $m \geq 2$. The choice $m = 2$ gives $b = 2$, which is rational, so m must avoid the powers of 2; the first choice left is $m = 3$.

Proposition 8.7.2. A Witness for an Irrational Power

Optional: nothing later depends on this.

The number $\log_2 9$ is irrational. Consequently $a := \sqrt{2}$ and $b := \log_2 9$ are irrational numbers with $a^b = 3$.

Idea. The irrationality goes as for $\sqrt{2}$: suppose $\log_2 9$ is a fraction p/q, clear the denominator, and look at parity. Here the clash is quicker than for $\sqrt{2}$: a power of 2 would equal a power of 9, and one is even while the other is odd. The positivity of $\log_2 9$ is what makes the power of 2 even.

Proof. Suppose that $\log_2 9 = p/q$ for integers p and q with $q \neq 0$; replacing p, q by $-p$, $-q$ if necessary, we may take $q \geq 1$. Since $2^0 = 1 < 9$ and $t \mapsto 2^t$ is increasing, $\log_2 9 > 0$, so $p > 0$, that is, $p \geq 1$. From $2^{p/q} = 9$, raising both sides to the power q gives

$$2^p = 9^q.$$

The left side is even, since $2^p = 2 \cdot 2^{p-1}$ with $p - 1 \geq 0$. The right side is odd, being a product of q odd factors. No integer is both even and odd, by (E2) of Section 2.1, so this is a contradiction. Hence $\log_2 9$ is not rational [¬I].

For the consequence, $a = \sqrt{2}$ is irrational by Theorem 2.2.1, and $b = \log_2 9$ is irrational by what was just shown. Since $\log_2 9 = \log_2(3^2) = 2\log_2 3$,

$$a^b = \left(2^{1/2}\right)^{\log_2 9} = 2^{(\log_2 9)/2} = 2^{\log_2 3} = 3.$$

This proves the proposition.

■

This proof uses no classical rule. The irrationality of $\log_2 9$ is a negative statement, "no fraction equals $\log_2 9$", and its box closes by ¬I, like the proof of Theorem 2.2.1. The witness $(\sqrt{2}, \log_2 9)$ is written down. So Proposition 8.7.2 gives a second proof of Theorem 8.7.1, and this one names its witness. (The exponent laws for real exponents are school facts here.)

Quick check. Does the proof of Proposition 8.7.2 use a classical rule? Does it name its witness?

Solution. No classical rule: its box assumes $\log_2 9 = p/q$ and concludes the negation, by ¬I. And yes, the witness $(\sqrt{2}, \log_2 9)$ is written down.

8.7.2 What a proof provides

Chapter 4 gave the connectives a meaning through truth values: every letter is true or false, and a table computes the rest. That semantics already takes a side, since a row decides every letter. The intuitionist explains a statement instead by saying what counts as a **proof** of it. The resulting clauses are called the **BHK reading**, after Brouwer, Heyting, and Kolmogorov.

To know a statement is to have a proof of it, and a proof of a compound statement is built from proofs of its parts.

- There is **no** proof of $\bot$.
- A proof of $P \wedge Q$ is a pair: a proof of P and a proof of Q.
- A proof of $P \vee Q$ is a proof of P or a proof of Q, **with a label saying which**.
- A proof of $P \to Q$ is a method that takes **any** proof of P and returns a proof of Q.
- A proof of $\neg P$ is a method that takes any proof of P and returns a proof of $\bot$. So $\neg P$ is read as $P \to \bot$, exactly the reading that the table in Section 2.2, after Definition 2.2.1, found row by row.
- A proof of $\exists x\, \varphi$ is an object t **together with** a proof of $\varphi[t/x]$.
- A proof of $\forall x\, \varphi$ is a method that produces, for **every** object t, a proof of $\varphi[t/x]$.

In words: a disjunction must say which side, an existence claim must name the object, and an implication or a universal statement must come with a method that works on every input. A proof of a negation provides no object at all; it is a guarantee that no proof of the negated statement will ever arrive.

Each rule other than RAA builds exactly what its clause asks for. ∨I writes $P \vee Q$ from a line P, so it knows which side it used. ∃I writes $\exists x\, \varphi$ from an instance $\varphi[t/x]$, so the witness t is on the page. A →I box is a method: given the assumption, it produces the conclusion. RAA is the exception. Its box shows that $\neg\varphi$ leads to $\bot$, which is a proof of $\neg\neg\varphi$, and then it writes φ. Under the BHK reading a method for refuting every refutation of φ is not a proof of φ.

Excluded middle under BHK. A proof of $P \vee \neg P$ for **every** statement P would be a single method that decides each statement, says which side holds, and proves that side. Nobody has such a method. Take Goldbach's conjecture, that every even integer greater than 2 is a sum of two primes: nobody has proved it and nobody has refuted it, so nobody can produce a labeled proof of "Goldbach's conjecture or its negation". The proof of Theorem 8.7.1 is of this kind: it

splits on "s is rational or not" by LEM, and so never says which pair works.

Not every, and some not. Chapter 6 derived $\neg\forall x\,\varphi \vdash \exists x\,\neg\varphi$ (Proposition 6.2.1 (a)) with two RAA boxes, and noted that the derivation never says which x fails φ. The BHK reading says why that matters. A proof of the premise is a method that refutes every proof of $\forall x\,\varphi$; a proof of the conclusion must name an object. Nothing in the first produces the second, and the derivation does not try: its conclusion comes from RAA, not from an $\exists$I applied to an available instance. Classical mathematics accepts the step. A constructive mathematician accepts that conclusion only with an object in hand.

Quick check. Describe a BHK proof of $\neg(P \wedge \neg P)$. Does it decide whether P holds?

Solution. We need a method that takes any proof of $P \wedge \neg P$ and returns a proof of $\bot$. A proof of $P \wedge \neg P$ is a pair: a proof p of P and a method f turning proofs of P into proofs of $\bot$. Apply f to p. That is the whole method, and it never decides P: it only uses a pair that someone else supplies. (It is the BHK reading of a $\neg$I box on $P \wedge \neg P$ whose clash is $\neg$E on its two halves.)

8.7.3 One extra rule

Chapter 2 found that one rule stands apart. RAA belongs to no connective, and over the other primitive rules it derives $\neg\neg$E and LEM and is derived by each of them (Theorem 2.5.7). Chapter 4's optional Theorem 4.7.4 showed that the other propositional rules cannot derive it.

Recall what Section 2.5 established. Over the primitive rules other than RAA, the three rules RAA, $\neg\neg$E, and LEM each derive the other two, so adding any one of them gives the same derivable sequents. There is one classical principle, and it has three faces. Nothing needs re-proving here; what is new is the decision to leave the principle out and see what remains.

Call a derivation **intuitionistic** if no line is justified by RAA and no line uses a derived rule or a cited result whose own derivation uses RAA. **Intuitionistic logic** is the system of these derivations. Its rules are the introduction and elimination rules for $\wedge$, $\rightarrow$, and $\vee$ (Chapter 1); $\neg$I, $\neg$E, $\bot$E, $\leftrightarrow$I, and $\leftrightarrow$E (Chapter 2); the rules for $\forall$, $\exists$, and $=$ (Chapter 5); and repetition. **Classical logic**, the logic of Chapters 1–7, is the same list with RAA added. TAUT (Section 4.8) is not among these rules: it is licensed by Corollary 4.8.5, that is, by the completeness theorem, whose proof splits on atoms by LEM (Lemma 4.8.3), so a line justified by TAUT, on a formula or on a propositional skeleton (Section 6.3), uses RAA in the sense just defined.

This is why Chapter 2 kept $\bot$E as a primitive rule although RAA and repetition derive it (the remark in the second-reading fold after Theorem 2.5.7). Had $\bot$E been dropped, striking RAA would have taken it away too, and the remaining system would be weaker than the one intuitionists actually use: it could no longer derive $\bot \vdash P$. That is a claim about all derivations, and we state it without proof; nothing depends on it. With $\bot$E kept, the rules without RAA are exactly intuitionistic logic.

System	Rules	$\vdash P \vee \neg P$?
intuitionistic	the primitive rules other than RAA	not derivable (Theorem 4.7.4 for the propositional rules, and the optional Section 8.8 again, by stages of knowledge; the rules for $\forall$, $\exists$ and $=$ do not change this, stated without proof)
classical	the same rules and RAA	derivable (Proposition 2.5.3)

Every intuitionistic derivation is a derivation, so everything intuitionistic logic derives, classical logic derives too, and soundness applies to it: Theorem 4.7.2, and, for derivations whose lines contain quantifiers or equations, the soundness statement for predicate logic in Section 7.2. The difference is a list of sequents that only the classical system reaches. Chapters 2–7 flag the derivations that use RAA as they go, so the results on that list can be read off the text; Exercise B4 asks for them.

Quick check. Is the derivation of Proposition 2.5.1 (a), $P \vdash \neg\neg P$, intuitionistic? Is that of (b)?

Solution.

(a) Yes: it is one $\neg$I box on $\neg P$. (b) No: it is derived with RAA.

8.7.4 Excluded middle cannot be refuted

If intuitionists do not assert $P \vee \neg P$, perhaps they deny it. They do not, and they could not: its double negation needs no classical rule at all.

Proposition 8.7.3. Excluded Middle Cannot Be Refuted

Optional: nothing later depends on this.

For every statement P, $\vdash \neg\neg(P \vee \neg P)$, by a derivation that uses neither RAA nor $\bot$E.

Try it first (ten minutes): the goal is a negation. Open the box that the goal asks for, hunt for a clash, and note which rule closes each box you open.

Idea. An attempt that starts from $P \vee \neg P$ stalls at once: $\vee$I needs one side, and neither side is in hand. But the goal is not $P \vee \neg P$. It is $\neg(\neg(P \vee \neg P))$, a negation, so the goal-directed move of Section 2.2 opens a $\neg$I box on $\neg(P \vee \neg P)$ with the goal $\bot$. That assumption can clash only with $P \vee \neg P$, which $\vee$I builds from either side, so we manufacture a side as in Example (Manufacturing the other half of a clash). A box on P gives $P \vee \neg P$, a clash, and $\neg$I closes it with $\neg P$. Then $\vee$I builds $P \vee \neg P$ from $\neg P$, and that is the second clash. Every box ends in $\bot$ and closes by $\neg$I.

Before reading on: self-explain lines 5 and 6.

Proof. Exercise B1 (a) of Section 3.2 (Exercise (B1: Three theorems)) derived this already, inside a conjunction; here is that derivation again, read for what it does not use. The derivation is:

1	$\neg(P \vee \neg P)$	asm, for $\neg$I
2	$\quad P$	asm, for $\neg$I
3	$\quad P \vee \neg P$	$\vee$I 2
4	$\quad \bot$	$\neg$E 3, 1
5	$\neg P$	$\neg$I 2–4
6	$P \vee \neg P$	$\vee$I 5
7	$\bot$	$\neg$E 6, 1
8	$\neg\neg(P \vee \neg P)$	$\neg$I 1–7

The two boxes close by $\neg$I, at lines 5 and 8. The other rules used are $\vee$I (lines 3 and 6) and $\neg$E (lines 4 and 7). The derivation has no premises. This proves the proposition.

■

Set it beside the derivation of Proposition 2.5.3. Lines 1–7 are the same, line for line. Only line 8 differs: there RAA closed the outer box and wrote $P \vee \neg P$; here $\neg$I closes it and writes the negation of its assumption. So the whole classical content of Proposition 2.5.3 sits in its last line.

Second reading · *Non-example by minimal change*

Keep lines 1–7, and write $P \vee \neg P$ at line 8 with the justification $\neg$I 1–7. The box still checks. What fails is the conclusion of $\neg$I, which must be the negation of the box's assumption, $\neg\neg(P \vee \neg P)$. Writing $P \vee \neg P$ there needs RAA.

What the proposition says is this: **nobody can refute excluded middle**. A refutation of $P \vee \neg P$, put beside this derivation, would prove $\bot$ from no premises, and no derivation does that. (That last fact is soundness: Theorem 4.7.2, or the soundness statement for predicate logic in Section 7.2 when P or the derivation involves quantifiers or equality, since no valuation, and no structure and assignment, makes $\bot$ true.) The intuitionist does not claim that excluded middle fails. The intuitionist declines to **assert** it, because asserting it would mean having a proof of one side, and for many statements P, an open conjecture for instance, nobody has one.

The proposition under BHK. The derivation of Proposition 8.7.3 is such a method, and it decides nothing. Its conclusion asks for a method that takes any f, a method sending proofs of $P \vee \neg P$ to proofs of $\bot$, and returns a proof of $\bot$. Here is one. Given a proof p of P, label it "left" and apply f: that turns proofs of P into proofs of $\bot$, which is a proof g of $\neg P$ (lines

2–5). Now label g "right" and apply f once more (lines 6–7). At no point did the method need to know whether P holds.

Remark. The proposition is one instance of a general fact, due to Glivenko (1929) and stated here without proof: if a statement built from letters by the connectives is derivable with all the rules, then its double negation is derivable without RAA. Classical propositional logic survives inside intuitionistic logic, behind two negations. The fact fails once quantifiers are allowed; $\neg\neg\forall x\,(\varphi \vee \neg\varphi)$ is in general not derivable without RAA (also stated without proof).

Warning. Intuitionistic logic does not say that excluded middle is false. No instance $P \vee \neg P$ can be refuted, since $\neg\neg(P \vee \neg P)$ is derivable without RAA; dropping a rule is not adding its negation. (Brouwer's own analysis adds axioms about the real numbers under which the **universal** statement "every real number is 0 or not 0" is refuted, but no single instance is.)

Quick check. Does Proposition 8.7.3 show that $P \vee \neg P$ is derivable without RAA?

Solution. No. It derives $\neg\neg(P \vee \neg P)$; writing $P \vee \neg P$ after the same box needs RAA, and Chapter 4's optional theorem in Section 4.7 shows that for a letter p no derivation without RAA reaches $p \vee \neg p$, as does the optional Section 8.8.

8.7.5 Which laws survive without RAA

Chapters 2 and 6 flagged every law they derived with RAA. Here they are side by side, split into directions. The middle column was derived without RAA. For the right column, with letters p, q, **no** derivation without RAA exists. The optional Section 8.8 proves this for the propositional rows; for the quantifier rows it is stated without proof.

Law	Derived without RAA	Needs RAA
double negation (Proposition 2.5.1)	(a) $P \vdash \neg\neg P$	(b) $\neg\neg P \vdash P$
contraposition (Proposition 2.3.2, Proposition 2.5.2)	$P \to Q \vdash \neg Q \to \neg P$	$\neg Q \to \neg P \vdash P \to Q$
De Morgan (Proposition 2.5.5)	(a), (b), (c)	(d) $\neg(P \wedge Q) \vdash \neg P \vee \neg Q$
implication as a disjunction (Proposition 2.5.6)	(a) $\neg P \vee Q \vdash P \to Q$	(b) $P \to Q \vdash \neg P \vee Q$
excluded middle (Proposition 2.5.3)	$\vdash \neg\neg(P \vee \neg P)$ (Proposition 8.7.3)	$\vdash P \vee \neg P$
Peirce's law (Proposition 2.5.4)	—	$\vdash ((P \to Q) \to P) \to P$
negated implication (Proposition 6.2.2)	$\varphi \wedge \neg\psi \vdash \neg(\varphi \to \psi)$	$\neg(\varphi \to \psi) \vdash \varphi \wedge \neg\psi$

Law	Derived without RAA	Needs RAA
modus tollens, disjunctive syllogism, constructive dilemma (Proposition 2.3.1, Proposition 2.3.4, Example (Constructive dilemma))	all	—
quantifier negation (Proposition 6.2.1)	(b); (c) both ways	(a) $\neg\forall x\,\varphi \vdash \exists x\,\neg\varphi$
passage laws (Proposition 6.2.7)	eight of the ten directions	$\dashv$ in (d) and in (e)

For each propositional entry of the right column, the double negation of its conclusion is derivable without RAA from the same premises, by Glivenko's theorem in the remark above (stated without proof) applied to the implication from the premises' conjunction; Exercise B5 (a) and (b) derive two cases.

Quick check. An intuitionist has proved $\neg Q \to \neg P$. May she conclude $P \to Q$?

Solution. Not in general: that step is Proposition 2.5.2, in the right column of the table. She may when Q is itself a negation, by Section 2.5's fold "RAA is never needed to finish a negation", which Section 8.1 applies to the contrapositive.

8.7.6 A short history

The dispute has dates.

In 1908 L. E. J. Brouwer published "De onbetrouwbaarheid der logische principes" ("The unreliability of the logical principles") in the *Tijdschrift voor Wijsbegeerte*. His point was that excluded middle had been abstracted from reasoning about finite collections, where every case can in principle be checked, and then applied without justification to infinite ones, where it cannot. Over the next two decades he rebuilt parts of analysis on this view, and the movement took the name **intuitionism**.

In 1925 A. N. Kolmogorov showed that classical reasoning can be imitated inside a logic without excluded middle by putting a double negation in front of every subformula: the first **double-negation translation** (Exercise C2 below checks one step of it).

In July 1927 David Hilbert lectured in Hamburg on "Die Grundlagen der Mathematik", printed in 1928. He defended excluded middle in a much-repeated comparison: to take it from the mathematician would be like forbidding the astronomer the telescope or the boxer the use of his fists. The quarrel between his program and Brouwer's became known as the **Grundlagenstreit**, the dispute over foundations. It turned personal in 1928, when Hilbert had Brouwer removed from the editorial board of the *Mathematische Annalen*; the board was then dissolved and re-formed without him.

In 1929 V. Glivenko proved the theorem of the remark above: classical propositional logic sits inside intuitionistic logic behind a double negation. In 1930 Arend Heyting published "Die formalen Regeln der intuitionistischen Logik" in the *Sitzungsberichte der Preußischen Akademie*

der Wissenschaften. It wrote down rules for intuitionistic logic, in a format different from ours, and from then on the logic could be studied as a formal system like any other.

In 1933 Kurt Gödel published a double-negation translation of arithmetic, which shows that classical arithmetic is consistent if intuitionistic arithmetic is. Gerhard Gentzen had found an equivalent translation independently and withdrew his paper of 1933 when Gödel's appeared.

In 1967 Errett Bishop published *Foundations of Constructive Analysis*. It developed a large part of analysis without excluded middle, and also without Brouwer's philosophy and his special axioms about the real numbers. Most of what is now called constructive analysis descends from it.

↻ **Second reading** · *Another dispute: the conditional*

Excluded middle is not the only rule that has been questioned. Chapter 1 read $P \to Q$ as a promise that is broken only when P holds and Q fails, so a true Q makes $P \to Q$ true whatever P says, and $\vdash P \to (Q \to P)$. Some logicians find that too generous: an "if" should connect its two parts. Logics built on that view, called **relevance logics**, reject $P \to (Q \to P)$. This book keeps Chapter 1's reading, which is the one mathematics uses.

8.7.7 What analysis would lose

What does analysis look like without the rule? Here are four previews, stated informally and without proof. Classical analysis takes the first as an axiom and proves the other three; each is a place where classical reasoning, by the rule or by an axiom, does work that a constructive proof cannot.

- **The supremum.** Classical analysis takes as an axiom that every non-empty set of real numbers that is bounded above has a supremum. Constructively, a non-empty set A bounded above has a supremum **if and only if** it is **order-located**: for all real numbers $x < y$, either y is an upper bound of A or some element of A exceeds x (Bishop's form of the least upper bound principle). The classical axiom asserts the supremum without requiring anyone to make these decisions.
- **Monotone sequences.** E. Specker (1949) gave an increasing, bounded sequence of rational numbers that a computer program can list term by term, but whose limit is not a computable real number: no program can print its decimal digits. So the convergence of bounded monotone sequences, which classical analysis proves from that axiom, cannot come with a method for computing the limit.
- **The intermediate value theorem.** Let f be continuous on $[a, b]$ with $f(a) < 0 < f(b)$. Classically, one proves that $f(x) = 0$ for some $x \in [a, b]$. Constructively, what holds in general is the **approximate** version: for every $\varepsilon > 0$ some $x \in [a, b]$ has $|f(x)| < \varepsilon$. The exact version holds under extra hypotheses, for example when f is strictly increasing. A bisection needs the sign of f at a midpoint m, and finitely many digits of $f(m)$ cannot always tell whether $f(m)$ is negative, zero, or positive: the digits of a value equal to 0 begin like those of a tiny positive or negative one.
- **Bolzano–Weierstrass.** A bisection proof of the theorem says: one of the two halves of an interval contains infinitely many terms of the sequence; continue with that half. That step is LEM on "the left half contains infinitely many terms", a statement that nobody can decide in general, and the proof does not say which half.

8.7.8 Proofs as programs

The BHK clauses read like the description of a programming language. A proof of $P \wedge Q$ is a pair; a proof of $P \vee Q$ is a value tagged "left" or "right"; a proof of $P \to Q$ is a function; a proof of $\exists x\, \varphi$ is a pair of an object and a proof; a proof of $\forall x\, \varphi$ is a function that accepts any object. This match between propositions and types, and between proofs and programs, is the **Curry–Howard correspondence**. It is what a proof assistant such as Lean is built on: checking a proof is checking that a program has the right type.

↻ **Second reading** · *Proofs as programs, in Lean*

Intuitionistic logic fits this picture exactly. LEM does not, since no program decides every proposition. Lean adds classical logic by an axiom, `Classical.choice`, a form of the axiom of choice. Excluded middle, `Classical.em`, is then a **theorem**, derived from choice together with two extensionality principles by Diaconescu's theorem. Lean draws one line with care. A definition that uses classical logic to produce **data**, for instance a number chosen by deciding an arbitrary proposition, must be marked `noncomputable`, since no program computes it. A proof that uses `Classical.em` needs no mark, because proofs carry no data.

```
-- excluded middle is a theorem, proved from the axiom of choice
example (p : Prop) : Or p (Not p) := Classical.em p

-- a number defined by deciding an arbitrary proposition
open Classical in
noncomputable def indicator (p : Prop) : Nat := if p then 1 else 0
```

Warning. "Constructive" does not mean "computable in practice". A constructive proof provides a method, and the method may be hopeless. Deciding whether a whole number is prime by trying every smaller divisor is constructive, and for a number with a thousand digits no computer could finish it. Constructive mathematics asks whether a method exists, not whether it is fast.

8.7.9 The book's position

The main line of this book stays classical. The decision is made once, in the axioms for the real numbers, which include trichotomy: for real numbers a and b, exactly one of $a < b$, $a = b$, $b < a$ holds. Constructive analysis does not accept trichotomy for real numbers, because it would decide the sign of every real number. So a case split such as "$x \geq 0$ or $x < 0$" rests on an axiom, and the book does not flag it.

What the book does flag is a genuinely non-constructive step of the kind discussed above, such as the bisection step of Bolzano–Weierstrass, and every use of the axiom of choice. One clause says so, and nothing more. Such a step is never silent anyway: by Definition 8.3.2, a $\neg\neg$E or LEM step is written out as an explicit case split or an explicit removal of a double negation. A reader who wants to know where the classical rule enters a proof can always find the place.

Remark (Looking ahead). *Analysis on the Real Line* takes trichotomy as one of its axioms in Chapter 1, and flags in one clause each genuinely non-constructive step it makes.

8.7.10 Exercises

8.7.10.1 A. Check your understanding

Exercise. A1

Parts (a)–(d) are about Proposition 8.7.3 and its derivation; part (e) is about the proof of Theorem 8.7.1.

(a) Which rule closes each of the two boxes, and at which line?

(b) The outer box opens with a negation and ends in $\bot$, the shape of an RAA box. Why is no RAA used?

(c) Give the BHK reading of the conclusion $\neg\neg(P \vee \neg P)$.

(d) True or false, with a reason: for some statement P, intuitionistic logic derives $\neg(P \vee \neg P)$ from no premises.

(e) Which line of that proof uses LEM, and on which statement? Does the proof say which pair (a, b) works?

Solution.

(a) The inner box, lines 2–4, opens with P and closes by $\neg$I at line 5, giving $\neg P$. The outer box, lines 1–7, opens with $\neg(P \vee \neg P)$ and closes by $\neg$I at line 8.

(b) The rule that closes a box is decided by what is written after it, not by the box. Line 8 writes $\neg\neg(P \vee \neg P)$, the negation of the box's assumption $\neg(P \vee \neg P)$, and that is what $\neg$I writes. RAA would be needed only to write $P \vee \neg P$ after the same box.

(c) A method that takes any method f, sending proofs of $P \vee \neg P$ to proofs of $\bot$, and returns a proof of $\bot$. The derivation provides one: label a proof of P "left" and apply f, which gives a proof of $\neg P$; label that proof "right" and apply f again.

(d) False. Such a derivation would also be a derivation with all the rules. Together with the derivation of $P \vee \neg P$ from no premises (Proposition 2.5.3), $\neg$E would derive $\bot$ from no premises. No derivation does that: by Theorem 4.7.2, or by the soundness statement for predicate logic in Section 7.2 when P or the derivation involves quantifiers or equality, $\bot$ would be true under every valuation (in every structure under every assignment), and no valuation, and no structure and assignment, makes $\bot$ true.

(e) The first line, which splits on "$\sqrt{2}^{\sqrt{2}}$ is rational" by LEM. The proof does not say which pair works: it shows that $(\sqrt{2}, \sqrt{2})$ works in the first case and $(\sqrt{2}^{\sqrt{2}}, \sqrt{2})$ in the second, and does not decide which case holds.

Exercise. A2: Drill

Answer each in a line. (a)–(d) ask whether a derivation avoids RAA; (e), (f) and (i) read proofs under BHK; the rest apply the section to proofs and claims.

(a) A derivation's last line is P, by RAA from a box on $\neg P$. Is the derivation intuitionistic?

(b) A box opens with P and ends in $\bot$, and the next line is $\neg P$. Which rule? Is the step intuitionistic?

(c) Is $\neg\neg\neg P \vdash \neg P$ derivable without RAA?

(d) A classmate derives $\neg P \lor \neg Q$ from $\neg(P \land Q)$ and says no RAA was used. What do you check first?

(e) Under BHK, what does a proof of $(P \land Q) \to P$ do?

(f) Under BHK, what must a proof of $\exists n \in \mathbb{N}\ (n^2 > 50)$ hand over?

(g) "$x \geq 0$ or $x < 0$; we consider the two cases." What is this in logical form, and does the book flag it?

(h) "Suppose $\sqrt{3}$ is rational. ... a contradiction. Hence $\sqrt{3}$ is irrational." Is this intuitionistic?

(i) A proof establishes $\neg\neg\exists x\, \varphi$. What does it hand over under BHK?

(j) True or false: every intuitionistic derivation is a derivation of Chapters 1–7.

(k) May an intuitionist argue "not every element of A is positive, so some element of A is not positive"?

(l) Which line of the proof of Theorem 8.7.1 would an intuitionist reject, and would she reject its conclusion?

Solution.

(a) No: the last line is justified by RAA.

(b) ¬I; yes.

(c) Yes: Example (Three negations become one) derives it, and its derivation uses no RAA.

(d) Every line and every cited result. That direction of De Morgan's law is in the right column of the table in "Which laws survive without RAA", so some line is RAA, ¬¬E or LEM, or cites a result derived with one of them.

(e) It takes any pair, a proof of P and a proof of Q, and returns the first component.

(f) A natural number together with a proof of the instance: for example 8, with $8 \in \mathbb{N}$ and $8^2 = 64 > 50$.

(g) LEM in form, an instance of trichotomy for the real numbers. The book does not flag it: it rests on an axiom ("The book's position").

(h) Yes: the goal is a negation, "$\sqrt{3}$ is not rational", and the box on "$\sqrt{3}$ is rational" closes by ¬I.

(i) No object: only a method that turns any refutation of $\exists x\, \varphi$ into a proof of $\bot$.

(j) True. It uses some of the rules of Chapters 1–7 and no others, so it is a derivation, and soundness applies to it.

(k) Not in general: that step is Proposition 6.2.3 (a), the bounded form of Proposition 6.2.1 (a), derived with RAA and in the right column of the table. She asks for the element.

(l) The first line, the LEM split on "s is rational". Not the conclusion: the second proof in "Irrational powers, with and without a witness" proves it with a named witness, $(\sqrt{2}, \log_2 9)$, and no classical rule.

8.7.10.2 B. Practice

Exercise. B1: Double negation of double negation elimination

Derive $\vdash \neg\neg(\neg\neg P \to P)$ without RAA, and name the rule on every line.

Hint: the goal is a negation; compare the derivation of the proposition in this section.

Solution. The goal is a negation, so we open a $\neg$I box on $\neg(\neg\neg P \to P)$, with the goal $\bot$. The assumption clashes only with $\neg\neg P \to P$, an implication, which $\to$I builds from a box on $\neg\neg P$ that ends in P. The first way to have P is to assume it: inside a box on P, the implication holds by $\to$I with P repeated, and it clashes with the assumption, so $\neg P$ follows by $\neg$I. Now $\neg P$ clashes with any $\neg\neg P$, so a second box on $\neg\neg P$ reaches $\bot$, $\bot$E writes P, and the implication is built a second time.

1	$\neg(\neg\neg P \to P)$	asm, for $\neg$I
2	P	asm, for $\neg$I
3	$\neg\neg P$	asm, for $\to$I
4	P	R 2
5	$\neg\neg P \to P$	$\to$I 3–4
6	$\bot$	$\neg$E 5, 1
7	$\neg P$	$\neg$I 2–6
8	$\neg\neg P$	asm, for $\to$I
9	$\bot$	$\neg$E 7, 8
10	P	$\bot$E 9
11	$\neg\neg P \to P$	$\to$I 8–10
12	$\bot$	$\neg$E 11, 1
13	$\neg\neg(\neg\neg P \to P)$	$\neg$I 1–12

Lines 1, 2, 3, and 8 are assumptions. Line 4 repeats line 2 into the box of line 3, since $\to$I needs P as a line of that box itself. Line 9 is $\neg$E on $\neg P$ (line 7) and its negation $\neg\neg P$ (line 8). Every box closes by $\to$I or $\neg$I, so no RAA is used. Unlike the derivation of the proposition in this section, this one uses $\bot$E, at line 10. For a letter p, the sequent $\vdash \neg\neg p \to p$ itself has no derivation without RAA (the optional Section 8.8 proves it with a two-stage model); its double negation has one.

Exercise. B2: Reading proofs as methods

(a) Give the BHK reading of a proof of $P \to \neg\neg P$, and describe such a proof.

(b) The same for $(P \to Q) \to (\neg Q \to \neg P)$.

(c) The same for $\forall n \in \mathbb{N}\ \exists m \in \mathbb{N}\ (m > n)$, over the natural numbers.

(d) A proof of an existence claim $\exists x\, \varphi$ opens with "suppose that no x satisfies φ" and

reaches a contradiction. Without a classical rule, which of $\neg\neg\exists x\,\varphi$ and $\exists x\,\varphi$ does it prove? What does it provide under the BHK reading?

Solution.

(a) A method that takes any proof p of P and returns a proof of $\neg\neg P$, that is, a method that takes any proof f of $\neg P$ and returns a proof of $\bot$. The method: given p, and then given f, apply f to p. This is the BHK reading of Proposition 2.5.1 (a).

(b) A method that takes any method h (from proofs of P to proofs of Q) and returns a proof of $\neg Q \to \neg P$: a method that takes any g (from proofs of Q to proofs of $\bot$) and returns a proof of $\neg P$, that is, a method from proofs of P to proofs of $\bot$. The method: given h, g, and a proof p of P, compute $g(h(p))$. This is the BHK reading of Proposition 2.3.2, whose derivation uses no RAA.

(c) Unfolding the bounded quantifiers (Definition 5.1.1): a method that takes any object n and a proof that $n \in \mathbb{N}$, and returns an object m together with a proof of $m \in \mathbb{N} \wedge m > n$. The method: given n and the proof of $n \in \mathbb{N}$, output $m = n + 1$, with the proofs that $n + 1 \in \mathbb{N}$ and $n + 1 > n$.

(d) The box opens with $\neg\exists x\,\varphi$ and ends in $\bot$, so $\neg$I closes it with $\neg\neg\exists x\,\varphi$. Writing $\exists x\,\varphi$ instead is RAA (or $\neg\neg$E after $\neg$I), a classical step. Under the BHK reading the proof provides a method that takes any proof that nothing satisfies φ and returns a proof of $\bot$. It provides no object: a proof of $\exists x\,\varphi$ would have to name one.

Exercise. B3: Where the witness goes missing

In the derivation of Proposition 6.2.1 (a), which step makes the existence claim without a witness? Why does no witness appear, although line 4 is an $\exists$I?

Solution. Two lines are classical: line 6, an RAA that writes the instance $\varphi[x_0/x]$ after a box on its negation, and line 9, the RAA that writes the conclusion $\exists x\,\neg\varphi$. Line 6 produces no existence claim. Line 9 is the step that asserts existence: it closes the box on $\neg\exists x\,\neg\varphi$, which ends in $\bot$, and writes $\exists x\,\neg\varphi$. Closed by $\neg$I, the same box would give only $\neg\neg\exists x\,\neg\varphi$.

No witness appears because no $\exists$I produces the conclusion. Line 4 does derive $\exists x\,\neg\varphi$ from the instance $\neg\varphi[x_0/x]$, but that instance is the assumption of the box 3–5, used only to clash with line 2 and discharged at line 6. The name x_0 is the flagged name of the $\forall$I box 3–6: an arbitrary object, not a particular one. Line 4 is unavailable at line 9, and the conclusion there is justified by RAA from the box 2–8, which names nothing. Under the BHK reading, line 9 asserts "something fails φ" and hands over no object that fails it.

Exercise. B4: Census of the classical results

(a) Which labeled results of Chapter 2 (propositions and theorems) were derived with RAA, directly or through $\neg\neg$E, LEM, or a cited result that uses one of them? (b) The same question for Chapters 3 to 7, counting lemmas, corollaries, and the step TAUT of Definition 4.8.1 as well.

Solution.

(a) Seven: Proposition 2.5.1 (b); Proposition 2.5.2; Proposition 2.5.3; Proposition 2.5.5 (d); Proposition 2.5.6 (b), through LEM; Theorem 2.5.7, whose proof uses all three: RAA in ①, ¬¬E in ②, LEM in ③; and Proposition 2.5.4. Every other proposition and theorem of Chapter 2 uses no classical rule, including Theorem 2.2.1, whose box closes by ¬I.

(b) Chapter 3: the direction ($\Leftarrow$) of Proposition 3.2.2 (b), whose box closes by RAA. Chapter 4: Lemma 4.8.3, through LEM, and with it Theorem 4.8.4 and the direction ($\Leftarrow$) of Corollary 4.8.5; the step TAUT of Definition 4.8.1, licensed through Corollary 4.8.5; and, in the optional Section 4.10, Lemma 4.10.2, whose key step uses RAA through Proposition 3.2.2, with Theorem 4.10.5 and Theorem 4.10.6, as the fold after the proof of Theorem 4.10.5 says. Chapter 5: none. In Chapter 6: Proposition 6.2.1 (a); the direction $\vdash$ of Proposition 6.2.2; the direction $\vdash$ of Proposition 6.2.3 (a); the two directions $\dashv$ of Proposition 6.2.7, in (d) and in (e), both through LEM; Proposition 6.2.5, as proved, since its box assumes $a \neq b$ and concludes $a = b$; and Proposition 6.3.1, through Corollary 4.8.5. Chapter 7: none derived; the classical theorems of Section 7.6, first-order completeness and what rests on it, are stated without proof, so the question does not reach them.

The results of Sections 4.4, 4.5 and 4.9, such as the standard equivalences and monotonicity, are not on the list: they are statements about valuations and derive nothing, and their remarks mention RAA only as history.

Exercise. B5: Behind two negations, and one instance

Derive each without RAA, ¬¬E, LEM or TAUT, and without citing a result derived with them.

(a) $\neg(P \wedge Q) \vdash \neg\neg(\neg P \vee \neg Q)$.

(b) $P \to Q \vdash \neg\neg(\neg P \vee Q)$.

(c) $\vdash (P \vee \neg P) \to (\neg\neg P \to P)$.

(d) Which row of the table in "Which laws survive without RAA" does each of (a)–(c) concern?

Solution.

(a) The goal is a negation, so a ¬I box on $\neg(\neg P \vee \neg Q)$. That assumption clashes with any disjunction $\neg P \vee \neg Q$, so we manufacture a side, as in the derivation of $\neg\neg(P \vee \neg P)$ in this section, twice:

1	$\neg(P \wedge Q)$	
2	$\quad \neg(\neg P \vee \neg Q)$	asm, for $\neg$I
3	$\quad\quad P$	asm, for $\neg$I
4	$\quad\quad\quad Q$	asm, for $\neg$I
5	$\quad\quad\quad P \wedge Q$	$\wedge$I 3, 4
6	$\quad\quad\quad \bot$	$\neg$E 5, 1
7	$\quad\quad \neg Q$	$\neg$I 4–6
8	$\quad\quad \neg P \vee \neg Q$	$\vee$I 7
9	$\quad\quad \bot$	$\neg$E 8, 2
10	$\quad \neg P$	$\neg$I 3–9
11	$\quad \neg P \vee \neg Q$	$\vee$I 10
12	$\quad \bot$	$\neg$E 11, 2
13	$\neg\neg(\neg P \vee \neg Q)$	$\neg$I 2–12

(b) The same plan, with the premise supplying Q inside the box on P:

1	$P \to Q$	
2	$\quad \neg(\neg P \vee Q)$	asm, for $\neg$I
3	$\quad\quad P$	asm, for $\neg$I
4	$\quad\quad Q$	$\to$E 1, 3
5	$\quad\quad \neg P \vee Q$	$\vee$I 4
6	$\quad\quad \bot$	$\neg$E 5, 2
7	$\quad \neg P$	$\neg$I 3–6
8	$\quad \neg P \vee Q$	$\vee$I 7
9	$\quad \bot$	$\neg$E 8, 2
10	$\neg\neg(\neg P \vee Q)$	$\neg$I 2–9

(c) Two implications, so two $\to$I boxes; inside, the instance of excluded middle is used by cases:

1	$P \vee \neg P$	asm, for $\to$I
2	$\neg\neg P$	asm, for $\to$I
3	P	asm, for $\vee$E
4	P	R 3
5	$\neg P$	asm, for $\vee$E
6	$\bot$	$\neg$E 5, 2
7	P	$\bot$E 6
8	P	$\vee$E 1, 3–4, 5–7
9	$\neg\neg P \to P$	$\to$I 2–8
10	$(P \vee \neg P) \to (\neg\neg P \to P)$	$\to$I 1–9

Every box closes by $\neg$I, $\to$I or $\vee$E. No line uses RAA, $\neg\neg$E or LEM, and no line cites a result. In (c), line 6 is $\neg$E on $\neg P$ (line 5) and $\neg\neg P$ (line 2).

(d) (a) concerns De Morgan (d): it is Glivenko's double negation of that row's conclusion. (b) concerns implication as a disjunction (b), in the same way. (c) concerns double negation (b): one instance of excluded middle, taken as a hypothesis, gives double-negation elimination at that instance with no further classical step. (The optional Section 8.8's Exercise C1 shows that such a step need not run back: one instance of excluded middle gives Peirce's law for that letter, and that instance of Peirce's law does not give the instance of excluded middle.)

Marking scheme. (10 points) - 3: (a). - 3: (b). - 3: (c). - 1: (d). - For each derivation: 1 for the skeleton (the outer $\neg$I or $\to$I box and its last line); 2 for the inside, each clash named.

Fatal errors (0 for that derivation): a line justified by RAA, $\neg\neg$E or LEM; a line justified by TAUT (each sequent is a tautology, so one TAUT line would "derive" it, but TAUT rests on completeness, which splits on atoms by LEM); citing Proposition 2.5.5 (d), Proposition 2.5.6 (b) or Proposition 2.5.1 (b), each derived with RAA.

8.7.10.3 C. Going deeper

Exercise. C1: More witnesses

This section proved that $\log_2 9$ is irrational; here are variants.

(a) Show that $a = \sqrt{2}$ and $b = \log_2 25$ are irrational numbers with a^b rational.

(b) For which integers $m \geq 2$ are $a = \sqrt{2}$ and $b = \log_2(m^2)$ irrational numbers with a^b rational? Prove your answer.

Solution (Hint 1). Compute a^b first; it does not depend on whether b is rational. Then write $m = 2^k r$ with r odd.

Solution (Hint 2). Suppose $\log_2(m^2) = p/q$ with integers p and $q \geq 1$; then $2^p = m^{2q}$.

Solution.

(a) First, $a^b = (2^{1/2})^{\log_2 25} = 2^{(\log_2 25)/2} = 2^{\log_2 5} = 5$, since $\log_2 25 = 2\log_2 5$; this is rational. Next, $a = \sqrt{2}$ is irrational by Theorem 2.2.1. Finally, suppose that $\log_2 25 = p/q$ with integers p and $q \geq 1$. Since $25 > 1 = 2^0$ and $t \mapsto 2^t$ is increasing, $\log_2 25 > 0$, so $p \geq 1$. Raising $2^{p/q} = 25$ to the power q gives $2^p = 25^q$. The left side is even because $p \geq 1$, and the right side is odd, being a product of odd factors. No integer is both, so $\log_2 25$ is irrational.

(b) **Answer:** exactly the integers $m \geq 2$ that are not powers of 2.

For every $m \geq 2$, $a = \sqrt{2}$ is irrational and

$$a^b = \left(2^{1/2}\right)^{\log_2(m^2)} = 2^{\log_2 m} = m,$$

which is rational. So the question is only when $b = \log_2(m^2) = 2\log_2 m$ is irrational. Write $m = 2^k r$ with $k \in \mathbb{N}$ and r odd, $r \geq 1$.

If $r = 1$, then $m = 2^k$ with $k \geq 1$, and $b = 2k$ is rational: these m fail.

If $r \geq 3$, suppose that $b = p/q$ with integers p and $q \geq 1$. Since $m \geq 2 > 1$, $b > 0$, so $p \geq 1$. Raising $2^{p/q} = m^2$ to the power q gives

$$2^p = m^{2q} = 2^{2kq} r^{2q}.$$

Since $r^{2q} \geq 3 > 1$, the right side exceeds 2^{2kq}, so $p > 2kq$. Dividing by 2^{2kq} gives $r^{2q} = 2^{p-2kq}$ with $p - 2kq \geq 1$. The left side is odd, a product of odd factors, and the right side is even. No integer is both, so b is irrational: these m work. The smallest are $m = 3, 5, 6, 7$. The cases $m = 3$ and $m = 5$ are the example of this section and part (a).

Marking scheme. (10 points) - 2: (a) $a^b = 5$, with the exponent laws used. - 2: (a) $\log_2 25$ irrational, with $p \geq 1$ justified by $\log_2 25 > 0$ and the parity clash $2^p = 25^q$. - 2: (b) $a^b = m$ for every m, so only the rationality of b matters. - 1: (b) powers of 2 excluded, with $b = 2k$ rational. - 3: (b) for $m = 2^k r$, $r \geq 3$ odd, the clash $r^{2q} = 2^{p-2kq}$ with $p - 2kq \geq 1$ (or an equivalent comparison of the odd parts of 2^p and m^{2q}). Fatal errors (0 for the step): concluding irrationality of b from "m is irrational" or "$\log_2 m$ is not an integer"; a parity clash with p possibly 0 or negative left unaddressed.

Exercise. C2: Modus ponens survives Kolmogorov's translation

For statements built from letters by $\wedge$, $\vee$, $\to$, and $\neg$, define the translation K by putting $\neg\neg$ in front of every subformula:

$$KP = \neg\neg P \text{ for a letter } P, \qquad K(\varphi \circ \psi) = \neg\neg(K\varphi \circ K\psi), \qquad K(\neg\varphi) = \neg\neg\neg K\varphi,$$

where $\circ$ is any of $\wedge$, $\vee$, $\to$. Chapter 2's Exercise (B1: With and without the new rule, used later) (a) derives $\neg\neg P$, $\neg\neg(P \to Q) \vdash \neg\neg Q$ without RAA (in its printed solution).

(a) Show that every $K\psi$ has the form $\neg\neg X$ for some statement X.

(b) Deduce that $K(\varphi \to \psi)$, $K\varphi \vdash K\psi$ without RAA, for all statements φ and ψ.

Solution (Hint 1). Put $K\varphi$ and $K\psi$ in for P and Q in the Chapter 2 sequent. Which of its two premises is $K(\varphi \to \psi)$?

Solution (Hint 2). The conclusion you get has two negations too many. Look for a result of Section 2.2 that removes negations without $\neg\neg$E.

Solution.

(a) By the definition, each of the three clauses produces a statement whose first two symbols are $\neg\neg$: $KP = \neg\neg P$, $K(\varphi \circ \psi) = \neg\neg(K\varphi \circ K\psi)$, and $K(\neg\varphi) = \neg\neg(\neg K\varphi)$. So $K\psi = \neg\neg X$, with X the letter, $K\varphi \circ K\psi$, or $\neg K\varphi$, respectively. No induction is needed, since the outermost clause alone decides the first two symbols.

(b) Three steps, each without RAA.

① The derivation of Exercise (B1: With and without the new rule, used later) (a) uses no RAA, so its sequent may be used as a derived rule (Definition 2.3.1) with the statements $K\varphi$ and $K\psi$ in place of P and Q:

$$\neg\neg K\varphi,\ \neg\neg(K\varphi \to K\psi) \vdash \neg\neg K\psi.$$

Its second premise is $K(\varphi \to \psi)$ itself, by the clause for $\to$.

② The first premise follows from $K\varphi$ by $\neg\neg$I (Proposition 2.5.1 (a), derived without RAA): a box on $\neg K\varphi$ clashes at once with $K\varphi$.

③ By (a), $K\psi = \neg\neg X$, so the conclusion of ① is $\neg\neg\neg\neg X = \neg\neg\neg(\neg X)$. Example (Three negations become one) derives $\neg\neg\neg P \vdash \neg P$ without RAA; with $\neg X$ in place of P it gives $\neg\neg\neg(\neg X) \vdash \neg(\neg X)$, that is, $\neg\neg K\psi \vdash K\psi$.

Chaining ②, ①, and ③ gives $K(\varphi \to \psi),\ K\varphi \vdash K\psi$, and no step is RAA: the derived rules used were all derived without it. In words, modus ponens survives the translation intuitionistically. (This is the translation that puts $\neg\neg$ in front of every subformula. In the Gödel–Gentzen translation of the history above, $\varphi \to \psi$ translates without an outer $\neg\neg$, and modus ponens survives trivially.)

Marking scheme. (10 points) - 2: (a) every clause of K begins with $\neg\neg$, read off the outermost clause. - 3: (b) the Chapter 2 sequent instantiated with $K\varphi$, $K\psi$, and its second premise recognized as $K(\varphi \to \psi)$. - 2: (b) $\neg\neg K\varphi$ from $K\varphi$ by $\neg\neg$I, named as RAA-free. - 3: (b) $\neg\neg K\psi \vdash K\psi$ by the triple-negation result applied to $\neg X$, using (a). Fatal errors (0 for the step): removing the extra negations by $\neg\neg$E (a classical step); assuming $K\psi$ is a letter.

8.7.10.4 R. Review

Exercise. R1

A set A of real numbers **has a maximum** if some element m of A satisfies $a \le m$ for every $a \in A$. Write "A has no maximum" in positive form, with no $\neg$ in front of a quantifier, and prove that the interval $(0, 1)$ has no maximum.

Solution. This uses the method of Chapter 6, Section 6.2: negate quantifier by quantifier, keeping the bounds (Proposition 6.2.3; compare Example (Three negations from analysis)). In symbols, "A has a maximum" is $\exists m \in A\, \forall a \in A\, (a \leq m)$. Its negation becomes, step by step,

$$\begin{aligned} \neg\exists m \in A\, \forall a \in A\, (a \leq m) &\dashv\vdash \forall m \in A\, \neg\forall a \in A\, (a \leq m) \\ &\dashv\vdash \forall m \in A\, \exists a \in A\, \neg(a \leq m), \end{aligned}$$

by part (b) of Proposition 6.2.3, and then part (a) used under the $\forall$, as Proposition 6.2.4 of Section 6.2 allows. For real numbers $\neg(a \leq m)$ says $a > m$. So the positive form is

$$\forall m \in A\, \exists a \in A\, (a > m)\,:$$

every element of A is exceeded by some element of A.

Now let $A = (0,1)$ and let $m \in (0,1)$. Take $a \coloneqq (m+1)/2$, which depends on m. Then

$$a - m = \frac{1-m}{2} > 0 \quad \text{and} \quad 1 - a = \frac{1-m}{2} > 0,$$

since $m < 1$, so $m < a < 1$; with $m > 0$ this gives $0 < a < 1$. So $a \in (0,1)$ and $a > m$. Since $m \in (0,1)$ was arbitrary, the positive form holds for $(0,1)$, and by the interderivability above, $(0,1)$ has no maximum. Going back from the positive form needs no classical step, but not by the route through Proposition 6.2.4, which uses both directions of (a). Go directly instead. Fix m with $m \in A$, as the flagged name of a $\forall$I box. The positive form gives $\exists a \in A\, (a > m)$ by $\forall$E and $\to$E; inside an $\exists$E box, $a > m$ gives $\neg(a \leq m)$ by trichotomy, and $\exists$I gives $\exists a \in A\, \neg(a \leq m)$. The direction $\dashv$ of (a) gives $\neg\forall a \in A\, (a \leq m)$, and $\to$I and $\forall$I give $\forall m \in A\, \neg\forall a \in A\, (a \leq m)$. The direction $\dashv$ of (b) then gives $\neg\exists m \in A\, \forall a \in A\, (a \leq m)$. Both directions $\dashv$ are derived without RAA, so this half of the argument uses no classical step beyond the order axioms.

Exercise. R2

Give a Fitch derivation of $P \to (Q \vee R)$, $Q \to S$, $R \to S \vdash P \to S$, and write it as one sentence of prose.

Solution. This uses Chapter 1's rules: $\to$I (Definition 1.4.1), $\to$E (Definition 1.3.1), and $\vee$E (Definition 1.6.1). The goal is an implication, so a box on P with the goal S. Inside it, the first premise gives $Q \vee R$, which we use by cases; each case reaches S through one of the other premises.

1	$P \to (Q \vee R)$	
2	$Q \to S$	
3	$R \to S$	
4	$\quad P$	asm, for $\to$I
5	$\quad Q \vee R$	$\to$E 1, 4
6	$\quad\quad Q$	asm, for $\vee$E
7	$\quad\quad S$	$\to$E 2, 6
8	$\quad\quad R$	asm, for $\vee$E
9	$\quad\quad S$	$\to$E 3, 8
10	$\quad S$	$\vee$E 5, 6–7, 8–9
11	$P \to S$	$\to$I 4–10

Line 10 may cite line 5, since both lie in the box of line 4, which is still open; the case boxes end with S as lines of their own. In prose: "Suppose P; then Q or R holds by the first premise, and S follows in the case Q by the second premise and in the case R by the third, so P implies S." Each premise is cited where it is used, as Definition 8.3.2 requires for every $\to$E on a hypothesis; the first premise is also the reason the two cases exhaust. The closing "so P implies S" is the closing line of $\to$I, which Definition 8.3.2 allows to go unsaid.

Exercise. R3

Decide with a truth table whether $(P \to Q) \to Q \models P \vee Q$, reading P and Q as atoms. If the entailment holds, give a derivation of $(P \to Q) \to Q \vdash P \vee Q$.

Solution. This item reaches back within this volume, since there is no earlier one: it uses the tables of Chapter 4 and the derivations of Chapter 2.

The table. Work backwards from F, as in Example (Working backwards from F) and the entailments of Definition 4.5.3. A refuting row needs $P \vee Q$ false, so P and Q are both false. In that row $P \to Q$ is true, so the premise $(P \to Q) \to Q$ is "T $\to$ F", which is false. So no row makes the premise true and the conclusion false, and the entailment holds. The full table agrees:

P	Q	$P \to Q$	$(P \to Q) \to Q$	$P \vee Q$
T	T	T	T	T
T	F	F	T	T
F	T	T	T	T
F	F	T	F	F

The derivation. The goal is a disjunction with neither side in hand, and the premise gives Q only once $P \to Q$ is available. We are stuck, so we open an RAA box on $\neg(P \vee Q)$ (Definition 2.5.1). Inside it, $P \to Q$ is built by a box on P: $\vee$I gives $P \vee Q$, which clashes with the assumption, and $\bot$E writes Q. Then the premise gives Q, and $\vee$I clashes again.

1	$(P \to Q) \to Q$	
2	$\quad \neg(P \vee Q)$	asm, for RAA
3	$\quad\quad P$	asm, for $\to$I
4	$\quad\quad P \vee Q$	$\vee$I 3
5	$\quad\quad \bot$	$\neg$E 4, 2
6	$\quad\quad Q$	$\bot$E 5
7	$\quad P \to Q$	$\to$I 3–6
8	$\quad Q$	$\to$E 1, 7
9	$\quad P \vee Q$	$\vee$I 8
10	$\quad \bot$	$\neg$E 9, 2
11	$P \vee Q$	RAA 2–10

The one classical line is line 11. Closing the same box by $\neg$I instead gives $\neg\neg(P \vee Q)$, and then the derivation is intuitionistic: the premise yields the double negation of its classical conclusion, the pattern of Glivenko's theorem in the remark of this section.

No derivation of this sequent avoids RAA or one of its equivalents; the optional Section 8.8 proves it with a two-stage model. (With $\bot$ in place of Q, the sequent would turn $\neg\neg P$ into P without any further classical step, and Chapter 4's optional theorem in Section 4.7 shows that $\neg\neg P \vdash P$ needs RAA.)

Exercise. R4

Prove that for every integer n, $n^3 - n$ is divisible by 3.

Solution. The tool is proof by cases on the remainder of n on division by 3: every integer n is $3k$, $3k+1$, or $3k+2$ for some integer k. This is a school fact, of the same kind as "every integer is even or odd" in Chapter 2. It is $\vee$E with three cases (Definition 1.6.1, used twice), and in each case $\exists$I (Definition 5.5.1) with an explicit witness. First factor:

$$n^3 - n = (n-1)\,n\,(n+1).$$

Let n be an integer, and let k be an integer with $n = 3k$, $n = 3k+1$, or $n = 3k+2$.

Case $n = 3k$. Then $n^3 - n = 3 \cdot \big((3k-1)\,k\,(3k+1)\big)$.

Case $n = 3k+1$. Then $n^3 - n = (3k)(3k+1)(3k+2) = 3 \cdot \big(k\,(3k+1)(3k+2)\big)$.

Case $n = 3k+2$. Then $n^3 - n = (3k+1)(3k+2)(3k+3) = 3 \cdot \big((3k+1)(3k+2)(k+1)\big)$.

In each case $n^3 - n = 3j$ for the integer j in brackets, so $3 \mid n^3 - n$. Since n was arbitrary, this holds for every integer n.

Why the look-alikes are worse. A proof by contradiction ("suppose $3 \nmid n^3 - n \ldots$") would have to reach exactly this case analysis and then clash with its assumption at the last line:

a direct proof in costume (Example (A direct proof in costume)). Induction is closer, but it fits the statement badly: with base case $P(0)$ and the step from n to $n+1$ (Axiom 4.2.1) it proves the claim only for $n \in \mathbb{N}$, and the negative integers $n = -m$, $m \in \mathbb{Z}_{>0}$, need a second argument, for instance $(-m)^3 - (-m) = -(m^3 - m)$. The case split covers all integers at once.

8.8 Stages of Knowledge: Kripke Models

This section is optional, and nothing outside it depends on it. Section 8.7 named several classical laws that have no derivation without RAA. For two of them, excluded middle and double-negation elimination, Chapter 4's optional Theorem 4.7.4 gave a proof by a three-valued reading; for the rest, Section 8.7 states the fact and points here. Truth tables cannot show it: every row makes $p \vee \neg p$ true. Here is a reading that covers them all. A statement is not true or false once and for all, but *known* or *not yet known* at a stage of an inquiry that grows. The rules without RAA are sound for this reading, and RAA is not; two or three stages then suffice to show which laws need it.

8.8.1 Stages of knowledge

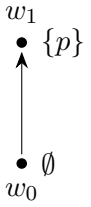
An inquiry passes through stages. At each stage some atomic statements are known: established, never to be withdrawn. From a stage, the inquiry may go on in several ways, to later stages where more is known. Nobody knows which way it will go. The atomic statements are the letters of Chapter 4's formulas (Definition 4.1.1), its atoms, and the formulas of this section are those formulas: letters, $\bot$, and the connectives $\neg$, $\wedge$, $\vee$, $\to$, $\leftrightarrow$.

Definition 8.8.1. Kripke Model

A **Kripke model** consists of:

- a finite non-empty set of **stages**;
- for each stage w, a set of stages **reachable** from w (written $w \le w'$), such that $w \le w$ for every stage, and $w \le w'$ and $w' \le w''$ give $w \le w''$;
- for each stage w, a set of letters **known at** w, such that a letter known at w is known at every stage reachable from w.

We draw a model as dots for the stages, with an arrow from a stage to each stage that comes directly after it. A stage is reachable from w when it can be reached by following arrows, none or several; so w reaches itself, and arrows are never drawn for that. Beside each dot stands the set of letters known there.



The model M_1: at w_0 nothing is known yet; at the later stage w_1, p is known.

8.8.2 Forcing

What a stage knows about a compound formula is computed from its parts, as a row computes a truth value, except that some connectives look ahead to later stages.

Definition 8.8.2. Forcing

Let w be a stage of a Kripke model. w **forces** a formula φ, written $w \Vdash \varphi$, as follows; $w \nVdash \varphi$ says that it does not. The definition is a recursion on φ (Theorem 4.3.4) that defines, for each formula, the set of stages that force it; clauses (K5) and (K6) need the sets of its parts, at other stages.

(K1) $w \Vdash p$, for a letter p, if p is known at w;
(K2) $w \Vdash \bot$ never;
(K3) $w \Vdash \varphi \wedge \psi$ if $w \Vdash \varphi$ and $w \Vdash \psi$;
(K4) $w \Vdash \varphi \vee \psi$ if $w \Vdash \varphi$ or $w \Vdash \psi$;
(K5) $w \Vdash \varphi \to \psi$ if every stage w' reachable from w with $w' \Vdash \varphi$ has $w' \Vdash \psi$;
(K6) $w \Vdash \neg\varphi$ if no stage reachable from w forces φ;
(K7) $w \Vdash \varphi \leftrightarrow \psi$ if $w \Vdash \varphi \to \psi$ and $w \Vdash \psi \to \varphi$.

A stage **refutes** a sequent $\Gamma \vdash \psi$ if it forces every formula of Γ and does not force ψ.

In words: a conjunction or a disjunction is decided at the stage itself, and a disjunction is forced only when one of its sides is: the inquiry must already know which. An implication and a negation look ahead. To force $\varphi \to \psi$ is to know that, however the inquiry goes on, ψ will be known whenever φ is. To force $\neg\varphi$ is to know that φ will never be known. Not forcing φ is weaker than forcing $\neg\varphi$.

A model with **one stage** is a row of a truth table. With one stage, the stages reachable from w are w alone, so (K5) says "φ is not forced, or ψ is", and (K6) says "φ is not forced": forcing becomes Chapter 4's truth (Definition 4.4.2). Kripke models are rows that can grow.

A stage that refutes $\Gamma \vdash \psi$ forces every member of Γ and not ψ, which is weaker than forcing $\neg\psi$. It is Section 1.1's counterexample set, read at a stage: where truth tables make $\neg\psi$ true, a stage only fails to know ψ.

Quick check. In a one-stage model in which p is not known, does the stage force $\neg p$? Does it force $p \vee \neg p$?

Solution. Yes to both. The only stage reachable is the stage itself, which does not force p, so (K6) gives $\neg p$, and (K4) gives $p \vee \neg p$. One stage is one row.

8.8.3 A first countermodel

Example. Two stages refute excluded middle

In M_1, decide whether w_0 forces p, $\neg p$, $p \vee \neg p$, $\neg\neg p$ and $\neg\neg p \to p$.

Solution.

- p is not known at w_0, so $w_0 \nVdash p$ (K1).
- w_1 is reachable from w_0 and forces p, so $w_0 \nVdash \neg p$ (K6).
- Hence $w_0 \nVdash p \vee \neg p$ (K4): w_0 refutes $\vdash p \vee \neg p$.
- Next, $\neg\neg p$. No stage reachable from w_0 forces $\neg p$: not w_0, as just shown, and not w_1, which forces p. So $w_0 \Vdash \neg\neg p$ (K6).
- So w_0 forces $\neg\neg p$ and not p: it refutes $\neg\neg p \vdash p$, and by (K5), applied with $w' = w_0$, $w_0 \nVdash \neg\neg p \to p$.

At w_0 the inquiry knows that p cannot be refuted, since p becomes known later on every way it can go, but it does not yet know p. Excluded middle asks it to know now which of p and $\neg p$ holds.

8.8.4 Knowledge persists

Letters stay known by the definition of a Kripke model. The same holds for every formula.

Lemma 8.8.1. Heredity

In every Kripke model, if $w \Vdash \varphi$ and $w \leq w'$, then $w' \Vdash \varphi$.

In words: what a stage forces, every later stage forces. The proof is a structural induction on φ (Theorem 4.3.1); it is in the fold.

↻ **Second reading** · *The proof of heredity*

Proof. We prove, by Theorem 4.3.1, that every formula φ has the property: for all stages $w \leq w'$ of every Kripke model, $w \Vdash \varphi$ gives $w' \Vdash \varphi$.

Letters. A letter forced at w is known at w (K1), hence known at w' by the condition on known letters in Definition 8.8.1, hence forced at w'.

$\bot$. No stage forces it (K2), so there is nothing to show.

$\wedge$ *and* $\vee$. If $w \Vdash \varphi \wedge \psi$, then w forces both parts (K3), so w' forces both by the induction hypothesis, and $w' \Vdash \varphi \wedge \psi$. For $\vee$, w forces one part (K4), and w' forces the same part by the induction hypothesis.

$\to$. Let $w \Vdash \varphi \to \psi$, and let w'' be reachable from w' with $w'' \Vdash \varphi$. Since $w \leq w'$ and $w' \leq w''$, transitivity gives $w \leq w''$, so $w'' \Vdash \psi$ by (K5) at w. Hence $w' \Vdash \varphi \to \psi$ by (K5). No induction hypothesis is used.

$\neg$. The same: every stage reachable from w' is reachable from w, and none of those forces φ (K6), so $w' \Vdash \neg\varphi$.

$\leftrightarrow$. By (K7), w forces the two implications, so w' forces both by the case $\rightarrow$, and $w' \Vdash \varphi \leftrightarrow \psi$.

Every clause of Definition 4.1.1 is covered, so every formula has the property. This proves the lemma.

■

Quick check. A stage w forces $\neg\varphi$, and w' is reachable from w. Why does w' force $\neg\varphi$, whatever φ is?

Solution. $w \Vdash \neg\varphi$ says that no stage reachable from w forces φ. The stages reachable from w' are among those, by transitivity, so none of them forces φ either. Nothing is assumed about φ itself.

8.8.5 The rules without RAA are sound for stages

Theorem 8.8.2. Soundness without RAA

Let Γ be a list of formulas and ψ a formula. If $\Gamma \vdash \psi$ by a derivation every line of which is a formula of Definition 4.1.1, in which no line uses RAA or TAUT, nor a derived rule or cited result whose derivation uses one of them or a line outside Definition 4.1.1, then every stage of every Kripke model that forces every formula of Γ forces ψ.

In words: a derivation without RAA *never leads from what a stage knows to what it does not know.*

TAUT (Section 4.8) is not among the rules here: it is licensed by completeness, whose proof uses LEM (Section 8.7), so a derivation with a TAUT line is not a derivation without RAA.

The proof is Chapter 4's proof of Theorem 4.7.2, Steps 1–4, with "every valuation making the context true" replaced by "every stage forcing the context". Each rule's case, in one line:

Rule	Why its case holds at a stage w that forces the context
R	the cited line itself, forced at w
$\wedge$I, $\wedge$E, $\vee$I, $\leftrightarrow$E	the clause (K3), (K4) or (K7) at w itself, with (K5) at $w' = w$ for $\leftrightarrow$E
$\rightarrow$E, $\neg$E	(K5) or (K6), applied at $w' = w$, which is reachable from w; for $\neg$E no stage forces both cited lines, so there is nothing to show
$\bot$E	no stage forces $\bot$, so there is nothing to show
$\vee$E	w forces one side, by (K4); the box of that side ends in its last line, forced at w
$\rightarrow$I, $\neg$I, $\leftrightarrow$I	a later stage w' forces the context too, **by heredity**, so the box applies at w'

Rule	Why its case holds at a stage w that forces the context
RAA	**fails**: its box on $\neg\varphi$ shows only that no stage reachable from w forces $\neg\varphi$, that is, $w \Vdash \neg\neg\varphi$; in M_1, w_0 forces $\neg\neg p$ and not p

Second reading · *The proof, case by case*

Proof. **Step 1 (reduction to primitive rules).** As in Step 1 of the proof of Theorem 4.7.2, a derived rule or cited result derived without RAA and TAUT abbreviates a pasted derivation in the propositional primitive rules (Chapters 1 and 2) other than RAA. So fix a derivation of ψ whose premises all belong to Γ and which uses only those rules, with lines $1, \ldots, n$; let φ_k be the formula on line k and Γ_k its context.

Step 2 (the induction). For $k \in \mathbb{N}$, let $S(k)$ be the statement: *if* $1 \leq k \leq n$, *then every stage of every Kripke model that forces every formula of* Γ_k *forces* φ_k. We prove $S(k)$ for every k by Theorem 4.2.2. Let $1 \leq k \leq n$ (otherwise there is nothing to prove), assume $S(j)$ for every $j < k$, and let w force every formula of Γ_k. Every line and box that line k cites lies before it (Definition 1.4.2), so the induction hypothesis holds for each cited line and for the last line of each cited box.

Step 3 (the cases). *Premises and assumptions:* φ_k belongs to Γ_k, so w forces it. *One-step rules:* a cited line i is available at line k, so $\Gamma_i \subseteq \Gamma_k$ (Lemma 4.7.1 (a)), and the hypothesis gives $w \Vdash \varphi_i$. Then R is the hypothesis for the cited line, and $\wedge$I, $\wedge$E and $\vee$I follow from (K3) and (K4) at w. For $\rightarrow$E, $w \Vdash \varphi \rightarrow \psi$ and $w \Vdash \varphi$, and (K5) with $w' = w$ gives $w \Vdash \psi$; $\leftrightarrow$E is the same after (K7). For $\neg$E, $w \Vdash \neg\varphi$ would say by (K6) that w, reachable from itself, does not force φ, against $w \Vdash \varphi$; so no stage forces Γ_k, and there is nothing to show. For $\bot$E, $w \Vdash \bot$ is impossible by (K2), and again there is nothing to show.

Box rules: if line k cites the box from line a to line b, then $\Gamma_b = \Gamma_k, \varphi_a$ (Lemma 4.7.1 (b)), so the hypothesis for line b says that every stage forcing Γ_k and φ_a forces φ_b. For $\rightarrow$I, line k is $\varphi \rightarrow \psi$. Let w' be reachable from w with $w' \Vdash \varphi$. By Lemma 8.8.1, w' forces every formula of Γ_k, so it forces Γ_k and φ, and hence ψ. So $w \Vdash \varphi \rightarrow \psi$ by (K5). $\neg$I is the same with $\psi = \bot$: a stage w' reachable from w that forced φ would force $\bot$, against (K2), so none does, and $w \Vdash \neg\varphi$ by (K6). $\leftrightarrow$I is two cases of $\rightarrow$I, joined by (K7). For $\vee$E, $w \Vdash \varphi \vee \psi$ by the hypothesis for the cited line, so w forces φ or ψ (K4); if φ, then w forces Γ_k and φ, and the first box gives $w \Vdash \chi$; if ψ, the second box does. For RAA the case fails, as the table says, and no derivation here uses it.

So $S(k)$ holds, and by Theorem 4.2.2 every line k satisfies it.

Step 4 (the last line). The last line lies in no box (Definition 1.4.2), so its context Γ_n consists of premises only, each of which belongs to Γ (Definition 1.2.3). A stage forcing every formula of Γ forces every formula of Γ_n, so by Steps 2 and 3 it forces $\varphi_n = \psi$. This proves the theorem.

■

Corollary 8.8.3. Kripke Countermodel

If some stage of some Kripke model refutes $\Gamma \vdash \psi$, then $\Gamma \vdash \psi$ has no derivation without RAA in the propositional rules whose lines are all formulas of Definition 4.1.1.

Proof. Suppose it had one. By Theorem 8.8.2 the stage would force ψ, and it does not; so it has none. ∎

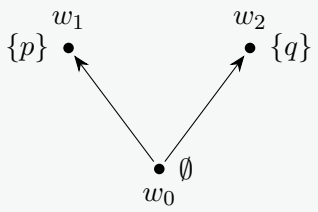
Quick check. Which case of the proof fails for RAA, and which stage of M_1 shows that it must?

Solution. RAA's own case: its box assumes $\neg\varphi$ and the rule writes φ. w_0 forces $\neg\neg p$ and not p. If RAA were sound for stages, the derivation of $\neg\neg p \vdash p$ (Proposition 2.5.1 (b)) would make w_0 force p.

8.8.6 Two branches

Example. A fork

In the model M_2 the inquiry can go two ways: one learns p, the other q.



Show that w_0 refutes $\vdash (p \to q) \vee (q \to p)$ and $\neg(p \wedge q) \vdash \neg p \vee \neg q$.

Solution.

- w_1 is reachable from w_0, forces p and does not force q, so $w_0 \nVdash p \to q$ (K5). Likewise w_2 gives $w_0 \nVdash q \to p$. By (K4), w_0 does not force the disjunction.
- No stage forces both p and q, so $w_0 \Vdash \neg(p \wedge q)$ (K6).
- w_1 forces p, so $w_0 \nVdash \neg p$; w_2 forces q, so $w_0 \nVdash \neg q$. Hence $w_0 \nVdash \neg p \vee \neg q$.
- By Corollary 8.8.3, neither $\vdash (p \to q) \vee (q \to p)$ nor $\neg(p \wedge q) \vdash \neg p \vee \neg q$ has a derivation without RAA.

Classically each is derivable (Section 2.5, and Corollary 4.8.5 for the first, a tautology). The branch is what defeats them. At w_0 the inquiry knows that p and q will never both be known, but not which one will fail.

8.8.7 What two or three stages settle

Here is the propositional right column of the table in Section 8.7, "Which laws survive without RAA", each sequent with a stage that refutes it. Two more models are described in words: M_3 has stages $w_0 \to w_1$, nothing known at w_0, and p and q known at w_1; M_4 has stages $w_0 \to w_1$, p known at both, and q known at w_1 only.

Sequent (letters p, q)	A refuting stage
$\vdash p \vee \neg p$; $\neg\neg p \vdash p$; $\vdash \neg\neg p \to p$	w_0 of M_1 (Example (Two stages refute excluded middle))
$\vdash ((p \to q) \to p) \to p$	w_0 of M_1: no stage forces $p \to q$, so both stages force $(p \to q) \to p$; $w_0 \nVdash p$
$\neg(p \to q) \vdash p \wedge \neg q$	w_0 of M_1: no stage forces $p \to q$, so $w_0 \Vdash \neg(p \to q)$; $w_0 \nVdash p$
$(p \to q) \to q \vdash p \vee q$ (Section 8.7, R3)	w_0 of M_1: as above, both stages force $(p \to q) \to q$; w_0 forces neither p nor q
$p \to q \vdash \neg p \vee q$	w_0 of M_3: $w_0 \Vdash p \to q$; $w_0 \nVdash \neg p$ since $w_1 \Vdash p$, and $w_0 \nVdash q$
$\neg q \to \neg p \vdash p \to q$	w_0 of M_4: no stage forces $\neg q$, so $w_0 \Vdash \neg q \to \neg p$; w_0 forces p and not q
$\vdash (p \to q) \vee (q \to p)$; $\neg(p \wedge q) \vdash \neg p \vee \neg q$	w_0 of M_2 (Example (A fork))

With Corollary 8.8.3, each row proves that the sequent has no derivation without RAA in the propositional rules; that lines outside Definition 4.1.1 and the rules for $\forall$, $\exists$ and $=$ do not change this is stated without proof (Section 8.7). This covers every propositional entry of Section 8.7's right column, and the claims stated without proof in its Exercises B1 and R3.

8.8.8 Searching for a countermodel

How were these models found? Here is the search for $(p \to q) \to q \vdash p \vee q$, run backwards from what must fail, in the manner of Section 4.6's search for a refuting row.

- Start with a stage w_0 that must not force $p \vee q$. So neither p nor q is known there.
- It must force $(p \to q) \to q$: every later stage that forces $p \to q$ must force q. The cheapest way is to make sure no stage forces $p \to q$.
- w_0 itself must then fail $p \to q$, so some stage reachable from it forces p and not q. Add w_1 after w_0, with p known.
- Check every clause at both stages: this is M_1.

In general:

- to refute a disjunction at w, refute both sides at w itself;
- to refute $\varphi \to \psi$ or $\neg\varphi$ at w, add a later stage that forces φ (and, for $\to$, not ψ);
- keep letters known once known.

The search is a guide, not a decision procedure as stated here.

Remark. Three facts, stated without proof.

- The converse of Theorem 8.8.2 holds: a sequent forced at every stage of every Kripke model that forces its premises has a derivation without RAA. This is completeness for

this semantics.

- From it, by placing a new first stage below two countermodels, follows the **disjunction property**: if $\vdash \varphi \vee \psi$ without RAA, then $\vdash \varphi$ or $\vdash \psi$ without RAA, the BHK clause for "or" made a theorem about derivations.
- Kripke models extend to the quantifiers, with a domain at each stage that may grow. In them $\neg\forall x\,\varphi \vdash \exists x\,\neg\varphi$ and the two passage directions of Section 8.7's table can be refuted.

Nothing in this book depends on these facts.

Warning (Not allowed: misreadings of forcing). Each sentence below is wrong; the tag names the clause it misreads.

- "$w \nVdash p$, so $w \Vdash \neg p$": (K6) looks at every stage reachable from w; in M_1, w_0 forces neither p nor $\neg p$.
- "$w \Vdash \varphi \to \psi$, because $w \nVdash \varphi$": (K5) looks at every later stage, not only at w.
- "$w \Vdash \varphi \vee \psi$, because every stage reachable from w, other than w itself, forces φ or ψ": (K4) is decided at w itself; in M_2, every stage reachable from w_0 other than w_0 forces p or q, and w_0 forces neither.
- "A picture with p known at w_0 and not at a later w_1": it is not a Kripke model, since a known letter stays known (Definition 8.8.1).

8.8.9 Exercises

8.8.9.1 A. Check your understanding

Exercise. A1

(a) In M_1, does w_1 force $\neg\neg p \to p$?

(b) Why is a one-stage Kripke model a row of a truth table?

(c) True or false: if w does not force φ, then w forces $\neg\varphi$.

(d) Which rule's case fails in the proof of Theorem 8.8.2, and how does M_1 show that it must fail?

Solution.

(a) Yes. The only stage reachable from w_1 is w_1, which forces p, so it forces the implication by (K5).

(b) With one stage, (K5) and (K6) look only at that stage, so the clauses become those of Definition 4.4.2, with "forced" for "true".

(c) False: w_0 of M_1, with $\varphi = p$, forces neither p nor $\neg p$.

(d) RAA's. w_0 forces $\neg\neg p$ and not p, but $\neg\neg p \vdash p$ is derived with RAA; if RAA's case held, w_0 would force p.

Exercise. A2: Drill

Decide each in a line, with the clause you use. M_1 and M_2 are the models of the text. M_5 is the chain $w_0 \to w_1 \to w_2$, with nothing known at w_0, p known at w_1, and p, q known at w_2.

(a) M_1: does w_1 force $\neg p$?
(b) M_1, with q known nowhere: does w_0 force $\neg q$?
(c) M_2: does w_0 force $p \vee q$?
(d) M_2: does w_1 force $\neg q$?
(e) M_5: does w_0 force $p \to q$?
(f) M_5: does w_1 force $\neg\neg q$?
(g) Is this a Kripke model: $w_0 \to w_1$, with p known at w_0 and not at w_1?
(h) In M_1, a student says "w_0 forces $\neg p$, because p is not known at w_0". What is wrong?
(i) M_2: does w_0 force $\neg p \to q$?
(j) M_5: does w_0 force $q \vee \neg q$?
(k) M_5: does w_0 force $q \to p$?
(l) Can a one-stage model refute $\neg\neg p \vdash p$?

Solution.

(a) No: w_1 is reachable from itself and forces p (K6).
(b) Yes: no stage reachable from w_0 forces q (K6).
(c) No: neither letter is known at w_0 (K4).
(d) Yes: the only stage reachable from w_1 is w_1, where q is not known (K6).
(e) No: w_1 is reachable, forces p, and does not force q (K5).
(f) Yes. The stages reachable from w_1 are w_1 and w_2; neither forces $\neg q$, since w_2 forces q and is reachable from both (K6). And $w_1 \nVdash q$, so $\neg\neg q$ is forced where q is not.
(g) No: a letter known at w_0 must be known at every stage reachable from it (Definition 8.8.1).
(h) (K6) needs p to fail at every reachable stage, and w_1 forces p.
(i) Yes (K5). Of the stages reachable from w_0, only w_2 forces $\neg p$ (w_0 and w_1 reach w_1, which knows p), and w_2 forces q.
(j) No: q is not known at w_0, and w_2, reachable from w_0, forces q, so $w_0 \nVdash \neg q$ (K4, K6).
(k) Yes: the only stage that forces q is w_2, which forces p (K5).
(l) No. At a single stage forcing is truth, and $\neg\neg p$ is true exactly when p is. A refutation needs a later stage.

8.8.9.2 B. Practice

Exercise. B1: A chain of three stages

Let M_6 be the chain $w_0 \to w_1 \to w_2$, with nothing known at w_0, q known at w_1, and p and q known at w_2.

(a) Decide whether w_0 forces $\neg\neg p$, $p \vee \neg p$, $q \to p$, $(q \to p) \vee (p \to q)$ and $\neg p \to \neg q$.

(b) Which sequents among $\vdash p \vee \neg p$, $\neg\neg p \vdash p$ and $\vdash (p \to q) \vee (q \to p)$ does w_0 refute?

(c) A **chain** is a Kripke model whose stages can be listed $u_0, \ldots, u_m$ so that $u_i \leq u_j$ exactly when $i \leq j$. Show that no stage of a chain refutes $\vdash (p \to q) \vee (q \to p)$.

Solution.

(a)
- $\neg\neg p$: yes. Every stage reachable from w_0 reaches w_2, which forces p, so none forces $\neg p$ (K6).
- $p \vee \neg p$: no. p is not known at w_0, and $w_0 \nVdash \neg p$ because of w_2 (K4).
- $q \to p$: no. w_1 forces q and not p (K5).
- $(q \to p) \vee (p \to q)$: yes. The only stage forcing p is w_2, which forces q, so $w_0 \Vdash p \to q$ (K5), and (K4) gives the disjunction.
- $\neg p \to \neg q$: yes. No stage reachable from w_0 forces $\neg p$, so the implication holds with nothing to check (K5).

(b) $\vdash p \vee \neg p$ and $\neg\neg p \vdash p$, since w_0 forces $\neg\neg p$ and not p. Not $\vdash (p \to q) \vee (q \to p)$, by (a).

(c) Suppose some stage w of a chain refuted it, that is, forced neither $p \to q$ nor $q \to p$. By (K5), some u reachable from w forces p and not q, and some v reachable from w forces q and not p. In a chain, $u \leq v$ or $v \leq u$. If $u \leq v$, then v forces p, since letters stay known, a contradiction. If $v \leq u$, then u forces q, a contradiction. So no stage of a chain refutes it, and the branch of M_2 is needed.

Marking scheme. (10 points) - 5: (a), one per formula, each with its clause. - 2: (b). - 3: (c), the two orders of u, v, each with persistence of letters.

Fatal errors (0 for the step): deciding (K5) or (K6) at w_0 alone; in (c), checking one chain only.

Exercise. B2: Derive without RAA, or break with a Kripke model

For each claim, either derive it without RAA and without TAUT (and without results derived with RAA), or give a stage of a Kripke model that refutes it.

(a) $\neg\neg(p \wedge q) \vdash \neg\neg p \wedge \neg\neg q$.

(b) $\neg\neg p \wedge \neg\neg q \vdash \neg\neg(p \wedge q)$.

(c) $\neg\neg(p \vee q) \vdash \neg\neg p \vee \neg\neg q$.

(d) $(p \to q) \to p \vdash \neg\neg p$.

(e) $\neg p \to p \vdash p$.

(f) $(p \vee q) \wedge \neg p \vdash q$.

(g) $(p \to q) \wedge (\neg p \to q) \vdash q$.

Solution. The verdict comes first in each part.

(a) **Derive.** Each conjunct is a negation, so each takes a ¬I box, whose clash builds $\neg(p \wedge q)$ against the premise:

1	$\neg\neg(p \wedge q)$	
2	$\quad\neg p$	asm, for $\neg$I
3	$\quad\quad p \wedge q$	asm, for $\neg$I
4	$\quad\quad p$	$\wedge$E 3
5	$\quad\quad \bot$	$\neg$E 4, 2
6	$\quad\neg(p \wedge q)$	$\neg$I 3–5
7	$\quad\bot$	$\neg$E 6, 1
8	$\neg\neg p$	$\neg$I 2–7
9	$\quad\neg q$	asm, for $\neg$I
10	$\quad\quad p \wedge q$	asm, for $\neg$I
11	$\quad\quad q$	$\wedge$E 10
12	$\quad\quad \bot$	$\neg$E 11, 9
13	$\quad\neg(p \wedge q)$	$\neg$I 10–12
14	$\quad\bot$	$\neg$E 13, 1
15	$\neg\neg q$	$\neg$I 9–14
16	$\neg\neg p \wedge \neg\neg q$	$\wedge$I 8, 15

(b) **Derive.** A $\neg$I box on $\neg(p \wedge q)$; inside it, $\neg q$ and then $\neg p$ are built, each clashing with a conjunct of the premise:

1	$\neg\neg p \wedge \neg\neg q$	
2	$\neg\neg p$	$\wedge$E 1
3	$\neg\neg q$	$\wedge$E 1
4	$\quad\neg(p \wedge q)$	asm, for $\neg$I
5	$\quad\quad p$	asm, for $\neg$I
6	$\quad\quad\quad q$	asm, for $\neg$I
7	$\quad\quad\quad p \wedge q$	$\wedge$I 5, 6
8	$\quad\quad\quad \bot$	$\neg$E 7, 4
9	$\quad\quad \neg q$	$\neg$I 6–8
10	$\quad\quad \bot$	$\neg$E 9, 3
11	$\quad\neg p$	$\neg$I 5–10
12	$\quad\bot$	$\neg$E 11, 2
13	$\neg\neg(p \wedge q)$	$\neg$I 4–12

(c) **Break**, at w_0 of M_2. Every stage reachable from w_0 is, or reaches, w_1 or w_2, each of which forces $p \vee q$, so none forces $\neg(p \vee q)$: $w_0 \Vdash \neg\neg(p \vee q)$. But $w_2 \Vdash \neg p$ (only w_2 is reachable from it, and p is not known there), so $w_0 \nVdash \neg\neg p$. Likewise $w_1 \Vdash \neg q$, so

$w_0 \nVdash \neg\neg q$, and w_0 does not force the disjunction (K4).

(d) **Derive.** A $\neg$I box on $\neg p$, in which $p \to q$ is built by $\bot$E and fed to the premise:

1	$(p \to q) \to p$	
2	$\quad \neg p$	asm, for $\neg$I
3	$\quad\quad p$	asm, for $\to$I
4	$\quad\quad \bot$	$\neg$E 3, 2
5	$\quad\quad q$	$\bot$E 4
6	$\quad p \to q$	$\to$I 3–5
7	$\quad p$	$\to$E 1, 6
8	$\quad \bot$	$\neg$E 7, 2
9	$\neg\neg p$	$\neg$I 2–8

(e) **Break**, at w_0 of M_1. No stage forces $\neg p$, so $w_0 \Vdash \neg p \to p$ with nothing to check, and $w_0 \nVdash p$.

(f) **Derive.** The disjunction is used by cases; the case p clashes with $\neg p$:

1	$(p \vee q) \wedge \neg p$	
2	$p \vee q$	$\wedge$E 1
3	$\neg p$	$\wedge$E 1
4	$\quad p$	asm, for $\vee$E
5	$\quad \bot$	$\neg$E 4, 3
6	$\quad q$	$\bot$E 5
7	$\quad q$	asm, for $\vee$E
8	$\quad q$	R 7
9	q	$\vee$E 2, 4–6, 7–8

(g) **Break**, at w_0 of M_3 (stages $w_0 \to w_1$, nothing known at w_0, p and q known at w_1). The only stage forcing p is w_1, which forces q, so $w_0 \Vdash p \to q$. No stage forces $\neg p$, since both reach w_1, which forces p; so $w_0 \Vdash \neg p \to q$ with nothing to check. Hence w_0 forces the conjunction (K3), and $w_0 \nVdash q$.

The look-alikes: (a) and (b) are converses, and both derivable. (c) is (a)'s look-alike with $\vee$ for $\wedge$, and it breaks. (e) is (d)'s look-alike and breaks: from $\neg p \to p$ one derives only $\neg\neg p$ (a box on $\neg p$, then $\to$E, $\neg$E and $\neg$I). (g) is (f)'s look-alike: a proof by cases on p and $\neg p$ with no disjunction to split, which needs $p \vee \neg p$. It is the pattern of Lemma 4.8.3 (Section 4.8) with no further premises, which is why completeness, and with it TAUT, rests on LEM.

Marking scheme. (10 points) - 6: the four derivations, 1.5 each. - 3: the three counter-models, 1 each, every clause checked. - 1: one correct remark on a look-alike pair.

A wrong verdict scores 0 for that item. Fatal errors (0 for that item): a derivation that uses RAA, $\neg\neg$E or LEM, or cites a result derived with one; a TAUT line (every claim is a tautology, so a TAUT line would "derive" each, the breakable (c), (e) and (g) included; TAUT rests on completeness); a "model" with a known letter lost at a later stage.

8.8.9.3 C. Going deeper

Exercise. C1: Peirce's law and one instance of excluded middle

(a) Derive $((p \to q) \to p) \to p$ from the premise $p \vee \neg p$ without RAA or TAUT.

(b) Derive without RAA or TAUT, or refute with a Kripke model: $\vdash (((p \to q) \to p) \to p) \to (p \vee \neg p)$.

(c) Compare (a), (b) and Exercise (B5: Behind two negations, and one instance) (c) with Theorem 2.5.7, in one sentence.

Solution.

(a) A $\to$I box on $(p \to q) \to p$, and inside it the premise used by cases. In the case $\neg p$, the implication $p \to q$ is built by $\bot$E and fed to the box's assumption:

1	$p \vee \neg p$	
2	$(p \to q) \to p$	asm, for $\to$I
3	p	asm, for $\vee$E
4	p	R 3
5	$\neg p$	asm, for $\vee$E
6	p	asm, for $\to$I
7	$\bot$	$\neg$E 6, 5
8	q	$\bot$E 7
9	$p \to q$	$\to$I 6–8
10	p	$\to$E 2, 9
11	p	$\vee$E 1, 3–4, 5–10
12	$((p \to q) \to p) \to p$	$\to$I 2–11

(b) **Refute**, at w_0 of the fork M_7: stages $w_0 \to w_1$ and $w_0 \to w_2$, with p known at w_1 only, and q known nowhere.

- At w_2 no stage reachable forces p, so $w_2 \Vdash p \to q$, and $w_2 \nVdash p$. So $w_2 \nVdash (p \to q) \to p$, and neither does w_0, which reaches w_2. w_1 forces p, hence $(p \to q) \to p$. The only stage forcing $(p \to q) \to p$ is w_1, and it forces p: so $w_0 \Vdash ((p \to q) \to p) \to p$ (K5).
- $w_0 \nVdash p$, and $w_0 \nVdash \neg p$ because of w_1. So $w_0 \nVdash p \vee \neg p$. Since w_0 is reachable from itself, forces $((p \to q) \to p) \to p$ and does not force $p \vee \neg p$, (K5) gives $w_0 \nVdash (((p \to q) \to p) \to p) \to (p \vee \neg p)$, so w_0 refutes the claim; by Corollary 8.8.3 it has no derivation without RAA.
- M_1 would not do: there no stage forces $p \to q$, so w_0 forces $(p \to q) \to p$ and not p, and so does not force Peirce's law; the only stage that does, w_1, forces $p \vee \neg p$. So the second branch is needed.

(c) Theorem 2.5.7 compares rules, which act on every statement at once. Instance by instance the three differ: one instance of excluded middle gives Peirce's law (a) and double-negation elimination (Exercise (B5: Behind two negations, and one instance) (c)) for that letter, while that instance of Peirce's law does not give the instance of

excluded middle back (b).

Marking scheme. (10 points) - 5: (a). The $\vee$E on the premise, with both cases inside the box of line 2 (2); the case $\neg p$ reaching p without RAA (2), for instance by building $p \to q$ with $\bot$E, or via $\neg\neg p$ as in B2 (d) (citing B2 (d) and Exercise (B5: Behind two negations, and one instance) (c) is allowed); the closing $\to$I (1). - 4: (b). A model whose stage forces $((p \to q) \to p) \to p$ and not $p \vee \neg p$, every clause checked. - 1: (c).

Fatal errors (0 for the step): using M_1 in (b), where the formula is forced; using RAA or TAUT in (a).

Exercise. C2: Challenge: three negations are one

Prove that in every Kripke model, for every stage w and every formula φ: $w \Vdash \neg\neg\neg\varphi$ if and only if $w \Vdash \neg\varphi$. Compare with Example (Three negations become one).

Solution. ($\Leftarrow$) Let $w \Vdash \neg\varphi$, and let w' be reachable from w. By Lemma 8.8.1, $w' \Vdash \neg\varphi$. Since w' is reachable from itself, $w' \nVdash \neg\neg\varphi$ (K6). So no stage reachable from w forces $\neg\neg\varphi$, and $w \Vdash \neg\neg\neg\varphi$.

($\Rightarrow$) Let $w \Vdash \neg\neg\neg\varphi$, and suppose some w' reachable from w forces φ. Every w'' reachable from w' forces φ by Lemma 8.8.1, so no such w'' forces $\neg\varphi$ (K6, with w'' reachable from itself), and $w' \Vdash \neg\neg\varphi$. But w' is reachable from w, and $w \Vdash \neg\neg\neg\varphi$ says no such stage forces $\neg\neg\varphi$: a contradiction. So no stage reachable from w forces φ, and $w \Vdash \neg\varphi$.

The comparison: Example (Three negations become one) derives $\neg\neg\neg P \vdash \neg P$ without RAA, and $\neg P \vdash \neg\neg\neg P$ is Proposition 2.5.1 (a) for $\neg P$. By Theorem 8.8.2 both directions hold at every stage. The exercise proves the same fact directly, from the clauses.

Marking scheme. (10 points) - 4: ($\Leftarrow$), with heredity and reflexivity named. - 5: ($\Rightarrow$), with heredity of φ used to get $w' \Vdash \neg\neg\varphi$. - 1: the comparison. - Second route: both directions from Theorem 8.8.2, applied to the derivations without RAA of Example (Three negations become one) with φ for P and Proposition 2.5.1 (a) with $\neg\varphi$ for P, earn the 9 points for the two directions; the 1 point for the comparison then goes to naming where the direct proof uses heredity.

Fatal errors (0 for the step): "not forced, so the negation is forced"; heredity used without being cited.