

## 4.4 Valuations and Truth Tables

Sections 4.1 and 4.3 made formulas into objects: every formula is built from atoms and  $\perp$  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  $p, q, r, p_1, p_2, \dots$ , and we use Chapter 1's shorthand for them (outer parentheses dropped, the usual precedence,  $\rightarrow$  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  $\varphi$  and  $\psi$ . 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  $\varphi$  and writes  $\psi$ , 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:

| $\varphi$ | $\psi$ | $\varphi \wedge \psi$ |
|-----------|--------|-----------------------|
| T         | T      | T                     |
| T         | F      | F                     |
| F         | T      | F                     |
| F         | F      | F                     |

Here T stands for “true” and F for “false”.

*Contradiction.* The rule  $\perp$ E writes **any** statement from  $\perp$ . If  $\perp$  were ever true,  $\perp$ E would lead from a true line to every statement, false ones included. So  $\perp$  must be false, always.

*Negation.* The rule  $\neg$ E writes  $\perp$  from  $\varphi$  and  $\neg\varphi$ . Since  $\perp$  is false,  $\varphi$  and  $\neg\varphi$  are never both true: if  $\varphi$  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  $\varphi$  is false,  $\neg\varphi$  is true. Negation flips the value.

*Implication.* The rule  $\rightarrow$ E writes  $\psi$  from  $\varphi \rightarrow \psi$  and  $\varphi$ . So in the combination “ $\varphi$  true,  $\psi$  false”, the implication cannot be true: that is the one row where  $\varphi \rightarrow \psi$  is false. The other three rows are forced by derivations. Proposition 2.1.2 derives  $P \rightarrow Q$  from  $\neg P$ , so whenever  $P$  is false (and  $\neg P$  therefore true),  $P \rightarrow Q$  must be true, whatever  $Q$  is. And  $P \rightarrow Q$  is derivable from  $Q$  alone (Exercise (B2: An assumption that is not used) (a) of Section 1.4):

|   |                   |                          |
|---|-------------------|--------------------------|
| 1 | $Q$               |                          |
| 2 | $P$               | asm, for $\rightarrow$ I |
| 3 | $Q$               | R 1                      |
| 4 | $P \rightarrow Q$ | $\rightarrow$ I 2–3      |

so whenever  $Q$  is true,  $P \rightarrow Q$  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  $\rightarrow$ , “if  $P$  is false, then  $P \rightarrow Q$  is true”, are often presented as a convention to be swallowed. Here they are forced by  $\perp E$ , 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  $\rightarrow$ . 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  $\varphi$ , infer  $\varphi \diamond \psi$ ” and an elimination rule “from  $\varphi \diamond \psi$ , infer  $\psi$ ”, and read a table off them as above. The first rule needs  $\varphi \diamond \psi$  true whenever  $\varphi$  is; the second needs  $\psi$  true whenever  $\varphi \diamond \psi$  is. In a row with  $\varphi$  true and  $\psi$  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  $v(p) = T$ , **false under**  $v$  if  $v(p) = F$ .

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  $v(p) = T$ ,  $v(q) = F$  and  $v(r) = F$  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  $\varphi$  **under**  $v$  is defined by recursion on  $\varphi$ :

(V1) for an atom  $p$ ,  $v(p)$  is the value the valuation assigns to  $p$ ; and  $v(\perp) := F$ ;

(V2)  $v(\neg\varphi) := T$  if  $v(\varphi) = F$ , and  $v(\neg\varphi) := F$  if  $v(\varphi) = T$ ;

(V3)  $v(\varphi \wedge \psi) := T$  if  $v(\varphi) = T$  **and**  $v(\psi) = T$ ; otherwise  $F$ ;

(V4)  $v(\varphi \vee \psi) := T$  if  $v(\varphi) = T$  **or**  $v(\psi) = T$  (or both); otherwise  $F$ ;

(V5)  $v(\varphi \rightarrow \psi) := F$  if  $v(\varphi) = T$  and  $v(\psi) = F$ ; otherwise  $T$ ;

(V6)  $v(\varphi \leftrightarrow \psi) := T$  if  $v(\varphi) = v(\psi)$ ; otherwise  $F$ .

We say  $\varphi$  is **true under**  $v$ , or  $v$  **makes**  $\varphi$  true, if  $v(\varphi) = T$ ; otherwise  $\varphi$  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:

|           |               | $\varphi$ | $\psi$ | $\varphi \wedge \psi$ | $\varphi \vee \psi$ | $\varphi \rightarrow \psi$ | $\varphi \leftrightarrow \psi$ |
|-----------|---------------|-----------|--------|-----------------------|---------------------|----------------------------|--------------------------------|
| $\varphi$ | $\neg\varphi$ | T         | T      | T                     | T                   | T                          | T                              |
| T         | F             | T         | F      | F                     | T                   | F                          | F                              |
| F         | T             | 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  $\perp$  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  $\rightarrow$ , and the value of  $\perp$ , 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  $\perp$ , 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  $v(p) = \text{F}$ ,  $v(q) = \text{T}$  and  $v(r) = \text{F}$ . Compute  $v(\varphi)$  for

$$\varphi = (p \vee \neg q) \rightarrow (q \wedge r).$$

*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): $v(q) = \text{T}$   | F       |
| $p \vee \neg q$                            | (V4): both parts F        | F       |
| $q \wedge r$                               | (V3): $v(r) = \text{F}$   | F       |
| $(p \vee \neg q) \rightarrow (q \wedge r)$ | (V5): the hypothesis is F | T       |

So  $v(\varphi) = \text{T}$ . 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  $\perp$ , 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  $v(p) = \text{T}$  and  $v(q) = \text{F}$ , 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  $v(p) = \text{F}$ ,  $v(q) = \text{T}$ ,  $v(r) = \text{T}$ . What is  $v((p \vee q) \rightarrow \neg r)$ ?

*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  $v(p) = \text{T}$  and  $v(q) = \text{F}$ . What is  $v(\neg p \vee (q \leftrightarrow \perp))$ ?

*Solution.* T.  $\neg p$  is F by (V2).  $q \leftrightarrow \perp$  is T by (V6), since both sides are F ( $\perp$  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  $\varphi$ . At an atom  $p$  the game ends: the verifier wins if  $v(p) = \text{T}$ , the falsifier if  $v(p) = \text{F}$ . Run it once on  $\neg(p \wedge q)$ , with  $v(p) = \text{T}$  and  $v(q) = \text{F}$ . 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  $\text{atoms}(\varphi)$ , the set of atoms occurring in  $\varphi$ , is defined by recursion:  $\text{atoms}(p) = \{p\}$  for an atom  $p$ ,  $\text{atoms}(\perp)$  is empty,  $\text{atoms}(\neg\varphi) = \text{atoms}(\varphi)$ , and the atoms of  $\varphi \circ \psi$ , for each binary connective  $\circ$ , are those of  $\varphi$  together with those of  $\psi$ . 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  $\varphi$  be a formula, and let  $v$  and  $w$  be valuations that agree on  $\text{atoms}(\varphi)$ . Then  $v(\varphi) = w(\varphi)$ .

*Idea.* Structural induction on  $\varphi$ , with the property “for **all** valuations  $v, w$  that agree on the atoms of  $\varphi$ ,  $v(\varphi) = w(\varphi)$ ”. 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  $\text{atoms}(\varphi)$ ,  $v(\varphi) = w(\varphi)$ .

*Atoms.* Let  $p$  be an atom and let  $v, w$  agree on  $\text{atoms}(p) = \{p\}$ . Then  $v(p) = w(p)$ , which

is  $C(p)$ .

$\perp$ . By (V1),  $v(\perp) = F = w(\perp)$  for all valuations  $v, w$ . So  $C(\perp)$  holds.

*Negation.* Assume  $C(\varphi)$ , and let  $v, w$  agree on  $\text{atoms}(\neg\varphi)$ . Since  $\text{atoms}(\neg\varphi) = \text{atoms}(\varphi)$ , they agree on  $\text{atoms}(\varphi)$ , 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  $\text{atoms}(\varphi \wedge \psi)$ . Every atom of  $\varphi$  is an atom of  $\varphi \wedge \psi$ , so  $v$  and  $w$  agree on  $\text{atoms}(\varphi)$ , and  $v(\varphi) = w(\varphi)$  by  $C(\varphi)$ . In the same way  $v(\psi) = w(\psi)$  by  $C(\psi)$ . By (V3),  $v(\varphi \wedge \psi) = T$  exactly when  $v(\varphi) = v(\psi) = T$ , that is, exactly when  $w(\varphi) = w(\psi) = T$ , that is, exactly when  $w(\varphi \wedge \psi) = T$ . 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 $w$            |
| $\rightarrow$     | (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  $\varphi$ . 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  $\text{atoms}(\varphi \vee \psi)$ . As in the previous case,  $v(\varphi) = w(\varphi)$  and  $v(\psi) = w(\psi)$ . By (V4),  $v(\varphi \vee \psi) = T$  exactly when one of  $v(\varphi), v(\psi)$  is T, that is, when one of  $w(\varphi), w(\psi)$  is T, that is, when  $w(\varphi \vee \psi) = T$ . So  $C(\varphi \vee \psi)$ .

*Implication.* Assume  $C(\varphi)$  and  $C(\psi)$ , and let  $v, w$  agree on  $\text{atoms}(\varphi \rightarrow \psi)$ . Again  $v(\varphi) = w(\varphi)$  and  $v(\psi) = w(\psi)$ . By (V5),  $v(\varphi \rightarrow \psi) = F$  exactly when  $v(\varphi) = T$  and  $v(\psi) = F$ , that is, when  $w(\varphi) = T$  and  $w(\psi) = F$ , that is, when  $w(\varphi \rightarrow \psi) = F$ . So  $C(\varphi \rightarrow \psi)$ .

*Biconditional.* Assume  $C(\varphi)$  and  $C(\psi)$ , and let  $v, w$  agree on  $\text{atoms}(\varphi \leftrightarrow \psi)$ . Again  $v(\varphi) = w(\varphi)$  and  $v(\psi) = w(\psi)$ . By (V6),  $v(\varphi \leftrightarrow \psi) = \text{T}$  exactly when  $v(\varphi) = v(\psi)$ , that is, when  $w(\varphi) = w(\psi)$ , that is, when  $w(\varphi \leftrightarrow \psi) = \text{T}$ . 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  $\varphi$ : if  $\varphi$  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,  $\dots$ ,  $2^n$ . 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  $\varphi$  give  $\varphi$  the same value.

*Quick check.* Let  $\varphi = p \rightarrow (q \vee p)$ . 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:  $\varphi$  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  $\varphi$  be a formula and let  $p_1, \dots, p_n$  be a list of distinct atoms that **includes every atom of  $\varphi$** . A **row** is an assignment of a truth value to each of  $p_1, \dots, p_n$ . The **truth table** of  $\varphi$  (on  $p_1, \dots, p_n$ ) lists, for each of the  $2^n$  rows, the value  $v(\varphi)$  for a valuation  $v$  that **agrees with the row** on  $p_1, \dots, p_n$ . We say  $\varphi$  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  $p_1, \dots, p_n$  the row’s values and every other atom F. **Independence of the choice:** two valuations that agree with the same row agree on  $p_1, \dots, p_n$ , hence on  $\text{atoms}(\varphi)$ , so by Lemma 4.4.1 they give  $\varphi$  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  $p_1, \dots, p_n$  is allowed to contain atoms that do not occur in  $\varphi$ ; 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  $\varphi$  has no atoms at all, as  $\neg \perp$  or  $\perp \rightarrow \perp$  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  $\varphi$  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  $((p \rightarrow q) \wedge p) \rightarrow q$  on  $p, q$ , and read off from it what  $\rightarrow E$  does.

*Solution.* Columns for the two atoms, then for each subformula in turn:

| $p$ | $q$ | $p \rightarrow q$ | $(p \rightarrow q) \wedge p$ | $((p \rightarrow q) \wedge p) \rightarrow q$ |
|-----|-----|-------------------|------------------------------|----------------------------------------------|
| 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  $\rightarrow E$ : the only row in which both given lines  $p \rightarrow q$  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  $\varphi$  and  $\psi$  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  $(p \rightarrow q) \leftrightarrow (\neg q \rightarrow \neg p)$ .

*Solution.*

| $p$ | $q$ | $p \rightarrow q$ | $\neg q$ | $\neg p$ | $\neg q \rightarrow \neg p$ | $(p \rightarrow q) \leftrightarrow (\neg q \rightarrow \neg p)$ |
|-----|-----|-------------------|----------|----------|-----------------------------|-----------------------------------------------------------------|
| 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  $p \rightarrow q$  and  $\neg q \rightarrow \neg p$  are identical: each is F only in row 2. For  $\neg q \rightarrow \neg p$  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  $\neg Q \rightarrow \neg P$  from  $P \rightarrow Q$  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

$$\varphi = (p \rightarrow q) \rightarrow (p \rightarrow (q \wedge r))$$

is false: first with the full table on  $p, q, r$ , then without it.

*Solution. The full table.*

| $p$ | $q$ | $r$ | $p \rightarrow q$ | $q \wedge r$ | $p \rightarrow (q \wedge r)$ | $(p \rightarrow q) \rightarrow (p \rightarrow (q \wedge r))$ |
|-----|-----|-----|-------------------|--------------|------------------------------|--------------------------------------------------------------|
| 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  $\varphi$ , and then to whatever it forces.

1.  $\varphi$  is F exactly when  $p \rightarrow q$  is T and  $p \rightarrow (q \wedge r)$  is F.
2.  $p \rightarrow (q \wedge r)$  is F exactly when  $p$  is T and  $q \wedge r$  is F.
3. With  $p$  true,  $p \rightarrow q$  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  $\varphi$  false:  $p = T, q = T, r = F$ . 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  $(a_1, \dots, a_n)$  of truth values (each  $a_i$  is T or F) a truth value  $f(a_1, \dots, a_n)$ . A formula  $\varphi$  **expresses**  $f$  if

$$v(\varphi) = f(v(p_1), \dots, v(p_n)) \quad \text{for every valuation } v.$$

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  $T \mapsto F$ ,  $F \mapsto T$ ; 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 “ $\varphi$  because  $\psi$ ”:

| $\varphi$ | $\psi$ | $\varphi$ because $\psi$ |
|-----------|--------|--------------------------|
| 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  $\varphi$ ”:

| $\varphi$ | it was proved before 1900 that $\varphi$ |
|-----------|------------------------------------------|
| 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  $\varphi$  nor  $\psi$ ” 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  $p \rightarrow q$ , 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  $p = \text{F}, q = \text{F}$  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

- Make the truth tables of  $\neg(p \wedge \neg q)$  and of  $(p \vee q) \rightarrow p$ .
- Make the truth table of  $p \leftrightarrow \neg p$ . Is it true in any row?
- Compute the value of  $(p \rightarrow \perp) \vee (\perp \rightarrow p)$  under the all-F valuation, and under the all-T valuation.
- How many rows has the truth table of a formula on the atoms  $p, q, r, s$ ?
- True or false, with a reason: if  $v(p) \neq w(p)$ , then  $v(p \rightarrow (q \rightarrow p)) \neq w(p \rightarrow (q \rightarrow p))$ .

*Solution.*

(a)

| $p$ | $q$ | $\neg q$ | $p \wedge \neg q$ | $\neg(p \wedge \neg q)$ | $p$ | $q$ | $p \vee q$ | $(p \vee q) \rightarrow p$ |
|-----|-----|----------|-------------------|-------------------------|-----|-----|------------|----------------------------|
| T   | T   | F        | F                 | T                       | T   | T   | T          | T                          |
| T   | F   | T        | T                 | F                       | T   | F   | T          | F                          |
| F   | T   | F        | F                 | T                       | F   | T   | T          | F                          |
| F   | F   | T        | F                 | T                       | F   | F   | F          | T                          |

The first formula has the same column as  $p \rightarrow q$ : false only in row TF.

(b) In the row  $p = T$ ,  $\neg p$  is F, and in the row  $p = F$ ,  $\neg p$  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  $\perp \rightarrow p$  has a false hypothesis, since  $v(\perp) = F$  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  $p \rightarrow (q \rightarrow p)$  is true under every valuation: if  $v(p) = F$ , the hypothesis is false; if  $v(p) = T$ , the conclusion  $q \rightarrow p$  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.

- Let  $v(p) = T$  and  $v(q) = F$ . What is  $v(p \rightarrow q)$ ?
- Same  $v$ : what is  $v(\neg p \rightarrow q)$ ?
- Same  $v$ : what is  $v(\perp \leftrightarrow q)$ ?
- How many rows has a table on  $p, q, r, s, t$ ?
- Only  $v(p) = T$  is known. Which of  $p \vee q$ ,  $p \wedge q$ ,  $q \rightarrow p$ ,  $\neg p \rightarrow r$  already have a value, and which value?
- Give a row making  $(p \vee q) \rightarrow (p \wedge q)$  false.
- Is there a row making  $p \rightarrow \neg p$  true?
- Work backwards from F: in which rows is  $p \rightarrow (q \rightarrow r)$  false?
- Find every row in which  $(p \vee q) \wedge \neg r$  is true, forcing what you can first.
- Valuations  $v, w$  agree on  $p$  and differ on  $q$ . Must  $v((p \rightarrow p) \vee q) = w((p \rightarrow p) \vee q)$ ?
- In the game at  $\neg(p \wedge q)$ , with  $v(p) = F$  and  $v(q) = T$ : who chooses at the  $\wedge$ , what should be chosen, and who wins?
- Is “not both  $\varphi$  and  $\psi$ ” truth-functional? If so, give a formula that expresses it.

*Solution.*

- F: a true hypothesis and a false conclusion.
- T: the hypothesis  $\neg p$  is F.
- T: both sides are F.
- $2^5 = 32$ .
- $p \vee q$  is T (one true disjunct is enough);  $q \rightarrow p$  is T (true conclusion);  $\neg p \rightarrow r$  is T (false hypothesis).  $p \wedge q$  waits for  $q$ .

- (f)  $p = \text{T}, q = \text{F}$  (or  $p = \text{F}, q = \text{T}$ ).
- (g) Yes,  $p = \text{F}$ .
- (h) Only  $p = q = \text{T}, r = \text{F}$ .
- (i) A true conjunction forces both parts:  $\neg r$  true forces  $r = \text{F}$ . A true  $p \vee q$  leaves three choices, so the rows are  $p, q, r = \text{TTF}, \text{TFF}, \text{FTF}$ .
- (j) Yes, both are  $\text{T}$ , since  $p \rightarrow p$  is  $\text{T}$  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  $\text{F}$  only when both parts are  $\text{T}$ . 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  $\text{F}$  to find every row on its atoms in which it is false, or show that there is none.

- (a)  $((p \vee q) \wedge ((p \rightarrow r) \wedge (q \rightarrow r))) \rightarrow r$ .
- (b)  $(p \rightarrow (q \rightarrow r)) \rightarrow ((p \rightarrow q) \rightarrow (p \rightarrow r))$ .
- (c)  $((p \rightarrow r) \vee (q \rightarrow r)) \rightarrow ((p \vee q) \rightarrow r)$ .
- (d)  $((p \rightarrow q) \wedge (q \rightarrow r)) \rightarrow (r \rightarrow p)$ .

*Solution.* (a) Suppose the formula is  $\text{F}$ . Then  $r$  is  $\text{F}$ , and the hypothesis is  $\text{T}$ , so each of  $p \vee q, p \rightarrow r, q \rightarrow r$  is  $\text{T}$  (V3). With  $r$  false,  $p \rightarrow r$  true forces  $p = \text{F}$ , and  $q \rightarrow r$  true forces  $q = \text{F}$ . Then  $p \vee q$  is  $\text{F}$ , 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  $\text{F}$ . Then  $p \rightarrow (q \rightarrow r)$  is  $\text{T}$  and  $(p \rightarrow q) \rightarrow (p \rightarrow r)$  is  $\text{F}$ ; the latter forces  $p \rightarrow q$  to be  $\text{T}$  and  $p \rightarrow r$  to be  $\text{F}$ , and the last forces  $p = \text{T}, r = \text{F}$ . Then  $p \rightarrow q$  true with  $p$  true forces  $q = \text{T}$ . Now  $p \rightarrow (q \rightarrow r)$  has a true hypothesis and the conclusion  $q \rightarrow r$ , which is  $\text{F}$  (true  $q$ , false  $r$ ), so it is  $\text{F}$ , a contradiction. False in no row.

(c) Suppose it is  $\text{F}$ . Then  $(p \vee q) \rightarrow r$  is  $\text{F}$ :  $r = \text{F}$  and  $p \vee q$  is  $\text{T}$ . And  $(p \rightarrow r) \vee (q \rightarrow r)$  is  $\text{T}$ : with  $r$  false,  $p \rightarrow r$  is  $\text{T}$  exactly when  $p = \text{F}$ , and  $q \rightarrow r$  exactly when  $q = \text{F}$ ; so at least one of  $p, q$  is  $\text{F}$ . Together with  $p \vee q$  true: exactly one of  $p, q$  is  $\text{T}$ . This leaves a choice, so split. The formula is false in exactly two rows:  $p = \text{T}, q = \text{F}, r = \text{F}$  and  $p = \text{F}, q = \text{T}, r = \text{F}$ . Check the first:  $q \rightarrow r$  is  $\text{T}$ , so the hypothesis is  $\text{T}$ ;  $p \vee q$  is  $\text{T}$  and  $r$  is  $\text{F}$ , so the conclusion is  $\text{F}$ .

(d) Suppose it is  $\text{F}$ . Then  $r \rightarrow p$  is  $\text{F}$ :  $r = \text{T}, p = \text{F}$ . The hypothesis must be  $\text{T}$ :  $p \rightarrow q$  is  $\text{T}$  because  $p$  is  $\text{F}$ , and  $q \rightarrow r$  is  $\text{T}$  because  $r$  is  $\text{T}$ , whatever  $q$  is. Nothing forces  $q$ . So the formula is false in exactly two rows:  $p = \text{F}, r = \text{T}$ , 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  $\perp$  occurs in it; the immediate parts of a binary-built formula are again binary-built. Define by recursion (Theorem 4.3.4) the number  $\text{occ}(\varphi)$  of atom occurrences and the number  $\text{bin}(\varphi)$  of binary connectives:

$$\begin{aligned}\text{occ}(p) &= 1, & \text{occ}(\varphi \circ \psi) &= \text{occ}(\varphi) + \text{occ}(\psi), \\ \text{bin}(p) &= 0, & \text{bin}(\varphi \circ \psi) &= \text{bin}(\varphi) + \text{bin}(\psi) + 1,\end{aligned}$$

for every atom  $p$  and every binary connective  $\circ$ . (For  $\perp$  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  $\varphi$  has

$$\text{occ}(\varphi) = \text{bin}(\varphi) + 1.$$

The property is  $C(\varphi)$ : *if  $\varphi$  is binary-built, then  $\text{occ}(\varphi) = \text{bin}(\varphi) + 1$ .*

*$\perp$  and negation.* Neither  $\perp$  nor a formula  $\neg\varphi$  is binary-built, so  $C$  holds for them: its “if” is never met.

*Atoms.*  $\text{occ}(p) = 1 = 0 + 1 = \text{bin}(p) + 1$ .

*Conjunction.* Assume  $C(\varphi)$  and  $C(\psi)$ , and let  $\varphi \wedge \psi$  be binary-built. Then  $\varphi$  and  $\psi$  are binary-built, so

$$\begin{aligned}\text{occ}(\varphi \wedge \psi) &= \text{occ}(\varphi) + \text{occ}(\psi) = (\text{bin}(\varphi) + 1) + (\text{bin}(\psi) + 1) \\ &= (\text{bin}(\varphi) + \text{bin}(\psi) + 1) + 1 = \text{bin}(\varphi \wedge \psi) + 1.\end{aligned}$$

*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  $\varphi \rightarrow \psi$  be binary-built. Its immediate parts  $\varphi$  and  $\psi$  are binary-built, so  $\text{occ}(\varphi) = \text{bin}(\varphi) + 1$  and  $\text{occ}(\psi) = \text{bin}(\psi) + 1$ . Hence

$$\begin{aligned}\text{occ}(\varphi \rightarrow \psi) &= \text{occ}(\varphi) + \text{occ}(\psi) = (\text{bin}(\varphi) + 1) + (\text{bin}(\psi) + 1) \\ &= (\text{bin}(\varphi) + \text{bin}(\psi) + 1) + 1 = \text{bin}(\varphi \rightarrow \psi) + 1,\end{aligned}$$

which is  $C(\varphi \rightarrow \psi)$ . 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  $(p \rightarrow q) \wedge (p \vee r)$  has four atom occurrences ( $p$  counted twice) and three binary connectives.

**Exercise. B3: A table for a named pattern**

Make the truth table of  $((p \vee q) \wedge \neg p) \rightarrow q$ , and say which derived rule of Chapter 2 it corresponds to, as [Example \(The truth table of modus ponens\)](#) did for  $\rightarrow E$ .

*Solution.*

| $p$ | $q$ | $p \vee q$ | $\neg p$ | $(p \vee q) \wedge \neg p$ | $((p \vee q) \wedge \neg p) \rightarrow q$ |
|-----|-----|------------|----------|----------------------------|--------------------------------------------|
| 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  $p \rightarrow q$ , 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  $p = T, q = F$  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  $p \rightarrow \neg q$ ,  $q \vee r$ ,  $r \rightarrow p$  and  $\neg(p \wedge r)$ .

If  $r$  were T, claim 3 would force  $p = T$  (a true implication with a true hypothesis has a true conclusion), and then  $p \wedge r$  would be T, against claim 4. So  $r = F$ . Then claim 2 forces  $q = T$ , and claim 1 forces  $p = F$ : with  $p$  true, its conclusion  $\neg q$  would be false. Check: the row  $p = F, q = T, r = F$  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  $\varphi$ , let  $f_v(\varphi)$  be the number of occurrences of atoms in  $\varphi$  that are false under  $v$ , counted with repetition.

- (a) Prove by structural induction that every  $\leftrightarrow$ -built  $\varphi$  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  $\varphi$  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  $\varphi$  is  $\leftrightarrow$ -built, then for every valuation  $v$ :  $v(\varphi) = \text{T}$  iff  $f_v(\varphi)$  is even. The cases  $\perp$ ,  $\neg$  and the other three binary connectives are vacuous, since such formulas are not  $\leftrightarrow$ -built.

*Atoms.* If  $v(p) = \text{T}$ , then  $f_v(p) = 0$ , which is even; if  $v(p) = \text{F}$ , then  $f_v(p) = 1$ , which is odd. Either way,  $v(p) = \text{T}$  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),  $v(\varphi \leftrightarrow \psi) = \text{T}$  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  $v_{\text{T}}$  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  $\varphi$  is built from atoms by  $\wedge, \vee, \rightarrow$  only, then  $v_T(\varphi) = T$ ”. The formulas  $\perp, \neg\varphi$  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  $\varphi$  and  $\psi$  are built in this way too, so the hypothesis applies to them. (1) For an atom  $p$ ,  $v_T(p) = T$  by the definition of  $v_T$ . (2) If  $v_T(\varphi) = T$  and  $v_T(\psi) = T$ , then  $v_T(\varphi \wedge \psi) = T$  by (V3). (3) If  $v_T(\varphi) = T$  and  $v_T(\psi) = T$ , then  $v_T(\varphi \vee \psi) = T$  by (V4), since at least one part is true. (4) If  $v_T(\varphi) = T$  and  $v_T(\psi) = T$ , then  $v_T(\varphi \rightarrow \psi) = T$  by (V5), because  $v_T(\varphi) = F$  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  $v_T(\varphi) = T$ , and its reason states  $v_T(\varphi) = F$ . 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  $v_T(\varphi \rightarrow \psi) = T$  is correct, but not for the reason given.

(c)  $p \rightarrow q$ . Under  $v_T$ ,  $v_T(p) = T$ , so the stated reason “ $v_T(p) = F$ ” is false, although  $v_T(p \rightarrow q) = T$  holds.

(d) (4) If  $v_T(\varphi) = T$  and  $v_T(\psi) = T$ , then  $v_T(\varphi \rightarrow \psi) = T$  by (V5), because the conclusion  $\psi$  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  $\rightarrow$  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  $v_T(\varphi) = T$  of that very case.
- 2: a formula, such as  $p \rightarrow q$ , 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  $\perp, \neg, \leftrightarrow$ .

### 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  $\perp, \rightarrow$  and  $\leftrightarrow$  any values, they play no part) two statements,  $\text{Ver}(\varphi)$  (“the verifier can always win at  $\varphi$ ”) and  $\text{Fal}(\varphi)$  (“the falsifier can always win at  $\varphi$ ”):

| $\varphi$            | $\text{Ver}(\varphi)$ holds iff             | $\text{Fal}(\varphi)$ holds iff             |
|----------------------|---------------------------------------------|---------------------------------------------|
| an atom $p$          | $v(p) = \text{T}$                           | $v(p) = \text{F}$                           |
| $\neg\theta$         | $\text{Fal}(\theta)$                        | $\text{Ver}(\theta)$                        |
| $\theta \wedge \chi$ | $\text{Ver}(\theta)$ and $\text{Ver}(\chi)$ | $\text{Fal}(\theta)$ or $\text{Fal}(\chi)$  |
| $\theta \vee \chi$   | $\text{Ver}(\theta)$ or $\text{Ver}(\chi)$  | $\text{Fal}(\theta)$ and $\text{Fal}(\chi)$ |

- (a) Say in words why each clause matches the rules of the game.
- (b) Prove by structural induction that for every such formula  $\varphi$ ,  $\text{Ver}(\varphi)$  holds iff  $v(\varphi) = \text{T}$ , and  $\text{Fal}(\varphi)$  holds iff  $v(\varphi) = \text{F}$ .
- (c) Why must the two statements of (b) be proved together?

*Solution* (Hint 1). Put “if  $\varphi$  is built from  $\neg, \wedge, \vee$ ” inside the property.

*Solution* (Hint 2). The step for  $\neg$  needs the claim about  $\text{Fal}$  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  $\theta$ , and the other way round.

(b) The property is  $C(\varphi)$ : if  $\varphi$  is built from  $\neg, \wedge, \vee$ , then  $(\text{Ver}(\varphi) \text{ iff } v(\varphi) = \text{T})$  and  $(\text{Fal}(\varphi) \text{ iff } v(\varphi) = \text{F})$ . For  $\perp, \rightarrow$  and  $\leftrightarrow$  it is vacuous. For an atom both halves are the definition.

*Negation.* Assume  $C(\theta)$ . Then  $\text{Ver}(\neg\theta)$  iff  $\text{Fal}(\theta)$  iff  $v(\theta) = \text{F}$  iff  $v(\neg\theta) = \text{T}$ , by (V2). Likewise  $\text{Fal}(\neg\theta)$  iff  $\text{Ver}(\theta)$  iff  $v(\theta) = \text{T}$  iff  $v(\neg\theta) = \text{F}$ .

*Conjunction.* Assume  $C(\theta)$  and  $C(\chi)$ .  $\text{Ver}(\theta \wedge \chi)$  iff both  $\text{Ver}(\theta)$  and  $\text{Ver}(\chi)$ , iff both parts are true, iff  $v(\theta \wedge \chi) = \text{T}$ , by (V3).  $\text{Fal}(\theta \wedge \chi)$  iff  $\text{Fal}$  holds of one part, iff one part is false, iff  $v(\theta \wedge \chi) = \text{F}$ .

*Disjunction.* The same with (V4):  $\text{Ver}$  iff one part is true,  $\text{Fal}$  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  $\text{Ver}$  uses the hypothesis about  $\text{Fal}$  for the part, and the  $\neg$  step for  $\text{Fal}$  uses the one about  $\text{Ver}$ . 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.