Chapter 2
Propositional Logic

2.1 Propositions and Connectives

Traditionally, logic is said to be the art (or study) of reasoning; so in order to describe logic in this tradition, we have to know what "reasoning" is. According to some traditional views reasoning consists of the building of chains of linguistic entities by means of a certain relation "... follows from ...", a view which is good enough for our present purpose. The linguistic entities occurring in this kind of reasoning are taken to be *sentences*, i.e. entities that express a complete thought, or state of affairs. We call those sentences *declarative*. This means that, from the point of view of natural language, our class of acceptable linguistic objects is rather restricted.

Fortunately this class is wide enough when viewed from the mathematician's point of view. So far logic has been able to get along pretty well under this restriction. True, one cannot deal with questions, or imperative statements, but the role of these entities is negligible in pure mathematics. I must make an exception for performative statements, which play an important role in programming; think of instructions as "goto, if ... then, else ...", etc. For reasons given below, we will, however, leave them out of consideration.

The sentences we have in mind are of the kind "27 is a square number", "every positive integer is the sum of four squares", "there is only one empty set". A common feature of all those declarative sentences is the possibility of assigning them a truth value, *true* or *false*. We do not require the actual determination of the truth value in concrete cases, such as for instance Goldbach's conjecture or Riemann's hypothesis. It suffices that we can "in principle" assign a truth value.

Our so-called *two-valued* logic is based on the assumption that every sentence is either true or false; it is the cornerstone of the practice of truth tables.

Some sentences are minimal in the sense that there is no proper part which is also a sentence, e.g. $5 \in \{0, 1, 2, 5, 7\}$, or $2 + 2 = 5$; others can be taken apart into smaller parts, e.g. "c is rational or c is irrational" (where c is some constant). Conversely, we can build larger sentences from smaller ones by using *connectives*. We know many connectives in natural language; the following list is by no means meant to be exhaustive: *and, or, not, if ... then ..., but, since, as, for, although, neither ... nor*

D. van Dalen, *Logic and Structure*, Universitext, DOI 10.1007/978-1-4471-4558-5_2,

... . In ordinary discourse, and also in informal mathematics, one uses these connectives incessantly; however, in formal mathematics we will economize somewhat on the connectives we admit. This is mainly for reason of exactness. Compare, for example, the following two sentences: "π is irrational, but it is not algebraic", "Max is a Marxist, but he is not humorless". In the second statement we may discover a suggestion of some contrast, as if we should be surprised that Max is not humorless. In the first case such a surprise cannot be so easily imagined (unless, e.g. one has just read that almost all irrationals are algebraic); without changing the meaning one can transform this statement into "π is irrational and π is not algebraic". So why use (in a formal text) a formulation that carries vague, emotional undertones? For these and other reasons (e.g. of economy) we stick in logic to a limited number of connectives, in particular those that have shown themselves to be useful in the daily routine of formulating and proving.

Note, however, that even here ambiguities loom. Each of the connectives already has one or more meanings in natural language. We will give some examples:

1. John drove on and hit a pedestrian.
2. John hit a pedestrian and drove on.
3. If I open the window then we'll have fresh air.
4. If I open the window then $1 + 3 = 4$.
5. If $1 + 2 = 4$, then we'll have fresh air.
6. John is working or he is at home.
7. Euclid was a Greek or a mathematician.

From 1 and 2 we conclude that "and" may have an ordering function in time. Not so in mathematics; "π is irrational and 5 is positive" simply means that both parts are the case. Time just does not play a role in formal mathematics. We could not very well say "π was neither algebraic nor transcendent before 1882". What we would want to say is "before 1882 it was unknown whether π was algebraic or transcendent".

In examples 3–5 we consider the implication. Example 3 will be generally accepted, it displays a feature that we have come to accept as inherent to implication: there is a relation between the premise and conclusion. This feature is lacking in examples 4 and 5. Nonetheless we will allow cases such as 4 and 5 in mathematics. There are various reasons to do so. One is the consideration that meaning should be left out of syntactical considerations. Otherwise syntax would become unwieldy and we would run into an esoteric practice of exceptional cases. This general implication, in use in mathematics, is called *material implication*. Some other implications have been studied under the names of *strict implication, relevant implication*, etc.

Finally 6 and 7 demonstrate the use of "or". We tend to accept 6 and to reject 7. One mostly thinks of "or" as something exclusive. In 6 we more or less expect John not to work at home, while 7 is unusual in the sense that we as a rule do not use "or" when we could actually use "and". Also, we normally hesitate to use a disjunction if we already know which of the two parts is the case, e.g. "32 is a prime or 32 is not a prime" will be considered artificial (to say the least) by most of us, since we already know that 32 is not a prime. Yet mathematics freely uses such superfluous disjunctions, for example "$2 \geq 2$" (which stands for "$2 > 2$ or $2 = 2$").

In order to provide mathematics with a precise language we will create an artificial, formal language, which will lend itself to mathematical treatment. First we will define a language for propositional logic, i.e. the logic which deals only with *propositions* (sentences, statements). Later we will extend our treatment to a logic which also takes properties of individuals into account.

The process of *formalization* of propositional logic consists of two stages: (1) present a formal language, (2) specify a procedure for obtaining *valid* or *true* propositions.

We will first describe the language, using the technique of *inductive definitions*. The procedure is quite simple: *First* give the smallest propositions, which are not decomposable into smaller propositions; *next* describe how composite propositions are constructed out of already given propositions.

Definition 2.1.1 The language of propositional logic has an alphabet consisting of

(i) *proposition symbols*: $p_0, p_1, p_2, \ldots$,
(ii) *connectives*: $\wedge, \vee, \rightarrow, \neg, \leftrightarrow, \bot$,
(iii) *auxiliary symbols*: (,).

The connectives carry traditional names:

$\wedge$ – *and*	– *conjunction*
$\vee$ – *or*	– *disjunction*
$\rightarrow$ – *if ..., then ...*	– *implication*
$\neg$ – *not*	– *negation*
$\leftrightarrow$ – *iff*	– *equivalence, bi-implication*
$\bot$ – *falsity*	– *falsum, absurdum*

The proposition symbols and $\bot$ stand for the indecomposable propositions, which we call *atoms*, or *atomic propositions*.

Definition 2.1.2 The set *PROP* of propositions is the smallest set X with the properties

(i) $p_i \in X (i \in N), \bot \in X$,
(ii) $\varphi, \psi \in X \Rightarrow (\varphi \wedge \psi), (\varphi \vee \psi), (\varphi \rightarrow \psi), (\varphi \leftrightarrow \psi) \in X$,
(iii) $\varphi \in X \Rightarrow (\neg\varphi) \in X$.

The clauses describe exactly the possible ways of building propositions. In order to simplify clause (ii) we write $\varphi, \psi \in X \Rightarrow (\varphi \Box \psi) \in X$, where $\Box$ is one of the connectives $\wedge, \vee, \rightarrow, \leftrightarrow$.

A warning to the reader is in order here. We have used Greek letters φ, ψ in the definition; are they propositions? Clearly we did not intend them to be so, as we want only those strings of symbols obtained by combining symbols of the alphabet in a correct way. Evidently no Greek letters come in at all! The explanation is that φ and ψ are used as variables for propositions. Since we want to study logic, we must use a language in which to discuss it. As a rule this language is plain, everyday English.

We call the language used to discuss logic our *meta-language* and φ and ψ are *meta-variables* for propositions. We could do without meta-variables by handling (ii) and (iii) verbally: if two propositions are given, then a new proposition is obtained by placing the connective $\wedge$ between them and by adding brackets in front and at the end, etc. This verbal version should suffice to convince the reader of the advantage of the mathematical machinery.

Note that we have added a rather unusual connective, $\bot$. It is unusual in the sense that it does not connect anything. *Logical constant* would be a better name. For uniformity we stick to our present usage. $\bot$ is added for convenience; one could very well do without it, but it has certain advantages. One may note that there is something lacking, namely a symbol for the true proposition; we will indeed add another symbol, $\top$, as an abbreviation for the "true" proposition.

Examples

$$(p_7 \to p_0),\ ((\bot \vee p_{32}) \wedge (\neg p_2)) \in PROP,$$
$$p_1 \leftrightarrow p_7,\ \neg\neg\bot,\ ((\to \wedge \notin PROP.$$

It is easy to show that something belongs to *PROP* (just carry out the construction according to Definition 2.1.2); it is somewhat harder to show that something does not belong to *PROP*. We will do one example:

$$\neg\neg\bot \notin PROP.$$

Suppose $\neg\neg\bot \in X$ and X satisfies (i), (ii), (iii) of Definition 2.1.2. We claim that $Y = X - \{\neg\neg\bot\}$ also satisfies (i), (ii) and (iii). Since $\bot, p_i \in X$, also $\bot, p_i \in Y$. If $\varphi, \psi \in Y$, then $\varphi, \psi \in X$. Since X satisfies (ii) $(\varphi\Box\psi) \in X$. From the form of the expressions it is clear that $(\varphi\Box\psi) \neq \neg\neg\bot$ (look at the brackets), so $(\varphi\Box\psi) \in X - \{\neg\neg\bot\} = Y$. Likewise one shows that Y satisfies (iii). Hence X is not the smallest set satisfying (i), (ii) and (iii), so $\neg\neg\bot$ cannot belong to *PROP*.

Properties of propositions are established by an inductive procedure analogous to Definition 2.1.2: first deal with the atoms, and then go from the parts to the composite propositions. This is made precise in the following theorem.

Theorem 2.1.3 (Induction Principle) *Let A be a property, then $A(\varphi)$ holds for all $\varphi \in PROP$ if*

(i) $A(p_i)$, *for all i, and* $A(\bot)$,
(ii) $A(\varphi), A(\psi) \Rightarrow A((\varphi\Box\psi))$,
(iii) $A(\varphi) \Rightarrow A((\neg\varphi))$.

Proof Let $X = \{\varphi \in PROP \mid A(\varphi)\}$, then X satisfies (i), (ii) and (iii) of Definition 2.1.2. So $PROP \subseteq X$, i.e. for all $\varphi \in PROP$ $A(\varphi)$ holds. □

We call an application of Theorem 2.1.3 a *proof by induction on φ*. The reader will note an obvious similarity between the above theorem and the principle of complete induction in arithmetic.

The above procedure for obtaining all propositions, and for proving properties of propositions is elegant and perspicuous; there is another approach, however, which has its own advantages (in particular for coding): Consider propositions as the result of a linear step-by-step construction. For example $((\neg p_0) \to \bot)$ is constructed by assembling it from its basic parts by using previously constructed parts: $p_0 \ldots \bot \ldots (\neg p_0) \ldots ((\neg p_0) \to \bot)$. This is formalized as follows.

Definition 2.1.4 A sequence $\varphi_0, \ldots, \varphi_n$ is called a *formation sequence* of φ if $\varphi_n = \varphi$ and for all $i \le n$ φ_i is atomic, or

$$\varphi_i = (\varphi_j \Box \varphi_k) \quad \text{for certain } j, k < i, \quad \text{or}$$
$$\varphi_i = (\neg \varphi_j) \quad \text{for certain } j < i.$$

Observe that in this definition we are considering strings φ of symbols from the given alphabet; this mildly abuses our notational convention.

Examples $\bot, p_2, p_3, (\bot \vee p_2), (\neg(\bot \vee p_2)), (\neg p_3)$ and $p_3, (\neg p_3)$ are both formation sequences of $(\neg p_3)$. Note that formation sequences may contain "garbage".

We now give some trivial examples of proof by induction. In practice we actually only verify the clauses of the proof by induction and leave the conclusion to the reader.

1. *Each proposition has an even number of brackets.*

Proof

(i) Each atom has 0 brackets and 0 is even.
(ii) Suppose φ and ψ have $2n$, resp. $2m$ brackets, then $(\varphi \Box \psi)$ has $2(n + m + 1)$ brackets.
(iii) Suppose φ has $2n$ brackets, then $(\neg \varphi)$ has $2(n + 1)$ brackets. □

2. *Each proposition has a formation sequence.*

Proof

(i) If φ is an atom, then the sequence consisting of just φ is a formation sequence of φ.
(ii) Let $\varphi_0, \ldots, \varphi_n$ and $\psi_0, \ldots, \psi_m$ be formation sequences of φ and ψ, then one easily sees that $\varphi_0, \ldots, \varphi_n, \psi_0, \ldots, \psi_m, (\varphi_n \Box \psi_m)$ is a formation sequence of $(\varphi \Box \psi)$.
(iii) This is left to the reader. □

We can improve on 2.

Theorem 2.1.5 *PROP is the set of all expressions having formation sequences.*

Proof Let F be the set of all expressions (i.e. strings of symbols) having formation sequences. We have shown above that $PROP \subseteq F$.

Let φ have a formation sequence $\varphi_0, \dots, \varphi_n$, we show $\varphi \in PROP$ by induction on n.

$$n = 0 : \varphi = \varphi_0 \text{ and by definition } \varphi \text{ is atomic, so } \varphi \in PROP.$$

Suppose that all expressions with formation sequences of length $m < n$ are in $PROP$. By definition $\varphi_n = (\varphi_i \Box \varphi_j)$ for $i, j < n$, or $\varphi_n = (\neg \varphi_i)$ for $i < n$, or φ_n is atomic. In the first case φ_i and φ_j have formation sequences of length $i, j < n$, so by the induction hypothesis $\varphi_i, \varphi_j \in PROP$. As $PROP$ satisfies the clauses of Definition 2.1.2, also $(\varphi_i \Box \varphi_j) \in PROP$. Treat negation likewise. The atomic case is trivial. Conclusion $F \subseteq PROP$. □

Theorem 2.1.5 is in a sense a justification of the definition of formation sequence. It also enables us to establish properties of propositions by ordinary induction on the length of formation sequences.

In arithmetic one often defines functions by recursion, e.g. exponentiation is defined by $x^0 = 1$ and $x^{y+1} = x^y \cdot x$, or the factorial function by $0! = 1$ and $(x + 1)! = x! \cdot (x + 1)$.

The justification is rather immediate: each value is obtained by using the preceding values (for positive arguments). There is an analogous principle in our syntax.

Example The number $b(\varphi)$ of brackets of φ, can be defined as follows:

$$\begin{cases} b(\varphi) = 0 \quad \text{for } \varphi \text{ atomic,} \\ b((\varphi \Box \psi)) = b(\varphi) + b(\psi) + 2, \\ b((\neg \varphi)) = b(\varphi) + 2. \end{cases}$$

The value of $b(\varphi)$ can be computed by successively computing $b(\psi)$ for its subformulas ψ.

We can give this kind of definition for all sets that are defined by induction. The principle of "definition by recursion" takes the form of "there is a unique function such that ...". The reader should keep in mind that the basic idea is that one can "compute" the function value for a composition in a prescribed way from the function values of the composing parts.

The general principle behind this practice is laid down in the following theorem.

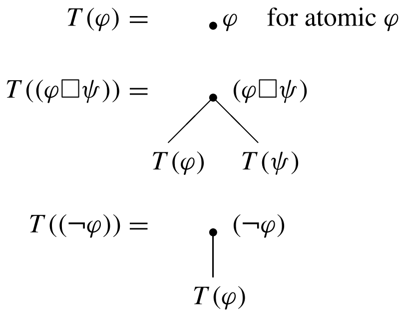
Theorem 2.1.6 (Definition by Recursion) *Let mappings $H_\Box : A^2 \to A$ and $H_\neg : A \to A$ be given and let H_{at} be a mapping from the set of atoms into A, then there exists exactly one mapping $F : PROP \to A$ such that*

$$\begin{cases} F(\varphi) = H_{at}(\varphi) \quad \textit{for } \varphi \textit{ atomic,} \\ F((\varphi \Box \psi)) = H_\Box(F(\varphi), F(\psi)), \\ F((\neg \varphi)) = H_\neg(F(\varphi)). \end{cases}$$

In concrete applications it is usually rather easily seen to be a correct principle. However, in general one has to prove the existence of a unique function satisfying the above equations. The proof is left as an exercise, cf. Exercise 11.

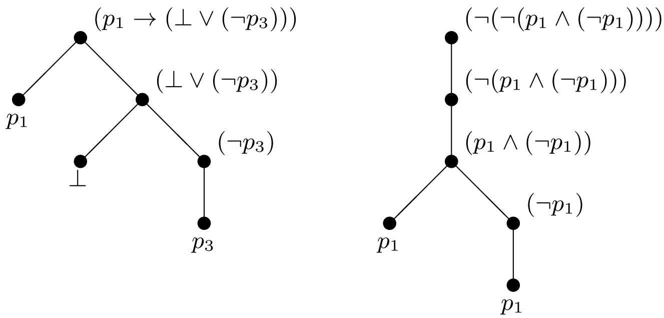
Here are some examples of definition by recursion.

1. The (parsing) *tree* of a proposition φ is defined by

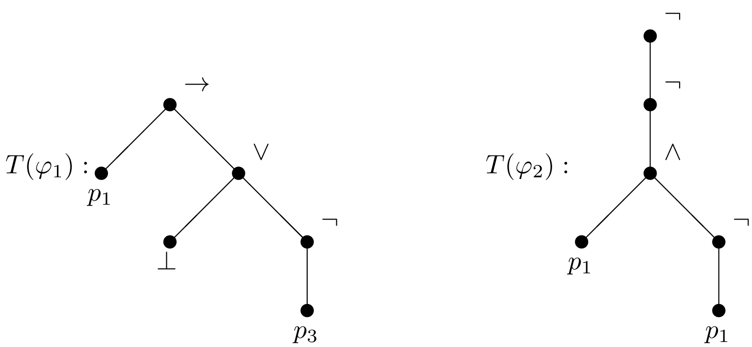


Examples

$T\big((p_1 \to (\bot \vee (\neg p_3)))\big);$ $\qquad$ $T\big(\neg(\neg(p_1 \wedge (\neg p_1)))\big)$



A simpler way to exhibit the trees consists of listing the atoms at the bottom, and indicating the connectives at the nodes.



2. The *rank* $r(\varphi)$ of a proposition φ is defined by

$$\begin{cases} r(\varphi) = 0 \quad \text{for atomic } \varphi, \\ r((\varphi \square \psi)) = \max(r(\varphi), r(\psi)) + 1, \\ r((\neg\varphi)) = r(\varphi) + 1. \end{cases}$$

We now use the technique of definition by recursion to define the notion of subformula.

Definition 2.1.7 The set of subformulas $Sub(\varphi)$ is given by

$$\begin{aligned} Sub(\varphi) &= \{\varphi\} \quad \text{for atomic } \varphi \\ Sub(\varphi_1 \square \varphi_2) &= Sub(\varphi_1) \cup Sub(\varphi_2) \cup \{\varphi_1 \square \varphi_2\} \\ Sub(\neg\varphi) &= Sub(\varphi) \cup \{\neg\varphi\}. \end{aligned}$$

We say that ψ *is a subformula of* φ if $\psi \in Sub(\varphi)$.

Examples p_2 is a subformula of $((p_7 \vee (\neg p_2)) \to p_1)$; $(p_1 \to \bot)$ is a subformula of $(((p_2 \vee (p_1 \wedge p_0)) \leftrightarrow (p_1 \to \bot))$.

Notational Convention In order to simplify our notation we will economize on brackets. We will always discard the outermost brackets and we will discard brackets in the case of negations. Furthermore we will use the convention that $\wedge$ and $\vee$ bind more strongly than $\to$ and $\leftrightarrow$ (cf. $\cdot$ and $+$ in arithmetic), and that $\neg$ binds more strongly than the other connectives.

Examples

$\neg\varphi \vee \varphi$	stands for $((\neg\varphi) \vee \varphi)$,
$\neg(\neg\neg\neg\varphi \wedge \bot)$	stands for $(\neg((\neg(\neg(\neg\varphi))) \wedge \bot))$,
$\varphi \vee \psi \to \varphi$	stands for $((\varphi \vee \psi) \to \varphi)$,
$\varphi \to \varphi \vee (\psi \to \chi)$	stands for $(\varphi \to (\varphi \vee (\psi \to \chi)))$.

Warning Note that those abbreviations are, properly speaking, not propositions.

In the proposition $(p_1 \to p_1)$ only one atom is used to define it; however it is used twice and it occurs at two places. For some purpose it is convenient to distinguish between *formulas* and *formula occurrences*. Now the definition of subformula does not tell us what an occurrence of φ in ψ is, we have to add some information. One way to indicate an occurrence of φ is to give its place in the tree of ψ, e.g. an occurrence of a formula in a given formula ψ is a pair (φ, k), where k is a node in the tree of ψ. One might even code k as a sequence of 0's and 1's, where we associate to each node the following sequence: $\langle\ \rangle$ (the empty sequence) to the top node, $\langle s_0, \ldots, s_{n-1}, 0\rangle$ to the left immediate descendant of the node with sequence $\langle s_0, \ldots, s_{n-1}\rangle$ and $\langle s_0, \ldots, s_{n-1}, 1\rangle$ to the second immediate descendant of it (if there is one). We will not be overly formal in handling occurrences of formulas (or symbols, for that matter), but it is important that it can be done.

The introduction of the rank function above is not a mere illustration of the "definition by recursion", it also allows us to prove facts about propositions by means of plain *complete induction* (or *mathematical induction*). We have, so to speak, reduced the tree structure to that of the straight line of natural numbers. Note that other "measures" will do just as well, e.g. the number of symbols. For completeness we will spell out the *Rank-Induction Principle*:

Theorem 2.1.8 (Rank-Induction Principle) *If for all* φ $[A(\psi)$ *for all* ψ *with rank less than* $r(\varphi)] \Rightarrow A(\varphi)$, *then* $A(\varphi)$ *holds for all* $\varphi \in PROP$.

Let us show that induction on φ and induction on the rank of φ are equivalent.[1]

First we introduce a convenient notation for the rank induction: write $\varphi \prec \psi$ ($\varphi \preceq \psi$) for $r(\varphi) < r(\psi)$ ($r(\varphi) \leq r(\psi)$). So $\forall \psi \preceq \varphi A(\psi)$ stands for "$A(\psi)$ holds for all ψ with rank at most $r(\varphi)$".

The *Rank-Induction Principle* now reads

$$\forall \varphi(\forall \psi \prec \varphi A(\psi) \Rightarrow A(\varphi)) \Rightarrow \forall \varphi A(\varphi)$$

We will now show that the rank-induction principle follows from the induction principle. Let

$$\forall \varphi(\forall \psi \prec \varphi A(\psi) \Rightarrow A(\varphi)) \tag{2.1}$$

be given. In order to show $\forall \varphi A(\varphi)$ we will indulge in a bit of induction loading. Put $B(\varphi) := \forall \psi \preceq \varphi A(\psi)$. Now show $\forall \varphi B(\varphi)$ by induction on φ.

1. For atomic φ $\forall \psi \prec \varphi A(\psi)$ is vacuously true, hence by (2.1) $A(\varphi)$ holds. Therefore $A(\psi)$ holds for all ψ with rank ≤ 0. So $B(\varphi)$.
2. $\varphi = \varphi_1 \Box \varphi_2$. Induction hypothesis: $B(\varphi_1), B(\varphi_2)$. Let ρ be any proposition with $r(\rho) = r(\varphi) = n + 1$ (for a suitable n). We have to show that ρ and all propositions with rank less than $n + 1$ have the property A. Since $r(\varphi) = \max(r(\varphi_1), r(\varphi_2)) + 1$, one of φ_1 and φ_2 has rank n—say φ_1. Now pick an arbitrary ψ with $r(\psi) \leq n$, then $\psi \preceq \varphi_1$. Therefore, by $B(\varphi_1)$, $A(\psi)$. This shows that $\forall \psi \prec \rho A(\psi)$, so by (2.1) $A(\rho)$ holds. This shows $B(\varphi)$.
3. $\varphi = \neg \varphi_1$. Similar argument.

An application of the induction principle yields $\forall \varphi B(\varphi)$, and as a consequence $\forall \varphi A(\varphi)$.

Conversely, the rank-induction principle implies the induction principle. We assume the premises of the induction principle. In order to apply the rank-induction principle we have to show (2.1). Now pick an arbitrary φ; there are three cases:

1. φ atomic. Then (2.1) holds trivially.
2. $\varphi = \varphi_1 \Box \varphi_2$. Then $\varphi_1, \varphi_2 \prec \varphi$ (see Exercise 6). Our assumption is $\forall \psi \prec \varphi A(\psi)$, so $A(\varphi_1)$ and $A(\varphi_2)$. Therefore $A(\varphi)$.
3. $\varphi = \neg \varphi_1$. Similar argument.

This establishes (2.1). So by rank induction we get $\forall \varphi A(\varphi)$.

---

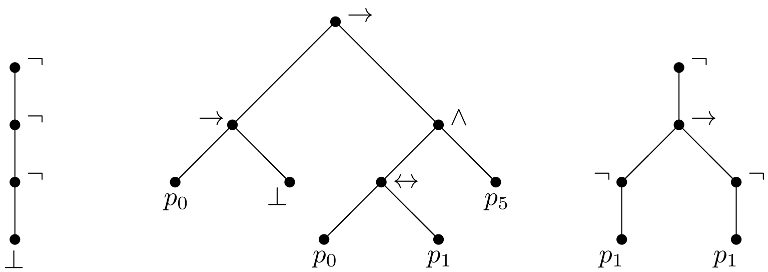
[1] The reader may skip this proof at first reading. He will do well to apply induction on rank naively.

Exercises

1. Give formation sequences of

$$(\neg p_2 \to (p_3 \vee (p_1 \leftrightarrow p_2))) \wedge \neg p_3,$$
$$(p_7 \to \neg \bot) \leftrightarrow ((p_4 \wedge \neg p_2) \to p_1),$$
$$(((p_1 \to p_2) \to p_1) \to p_2) \to p_1.$$

2. Show that $((\to \notin PROP$.
3. Show that the relation "is a subformula of" is transitive.
4. Let φ be a subformula of ψ. Show that φ occurs in each formation sequence of ψ.
5. If φ occurs in a shortest formation sequence of ψ then φ is a subformula of ψ.
6. Let r be the rank function.
   (a) Show that $r(\varphi) \leq$ number of occurrences of connectives of φ.
   (b) Give examples of φ such that $<$ or $=$ holds in (a).
   (c) Find the rank of the propositions in Exercise 1.
   (d) Show that $r(\varphi) < r(\psi)$ if φ is a proper subformula of ψ.
7. (a) Determine the trees of the propositions in Exercise 1.
   (b) Determine the propositions with the following trees.



8. Let $\#(T(\varphi))$ be the number of nodes of $T(\varphi)$. By the "number of connectives in φ" we mean the number of occurrences of connectives in φ. (In general $\#(A)$ stands for the number of elements of a (finite) set A.)
   (a) If φ does not contain $\bot$, show: number of connectives of φ + number of atoms of $\varphi \leq \#(T(\varphi))$.
   (b) $\#(\text{sub}(\varphi)) \leq \#(T(\varphi))$.
   (c) A branch of a tree is a maximal linearly ordered set.
   The length of a branch is the number of its nodes minus one. Show that $r(\varphi)$ is the length of a longest branch in $T(\varphi)$.
   (d) Let φ not contain $\bot$. Show: the number of connectives in φ + the number of atoms of $\varphi \leq 2^{r(\varphi)+1} - 1$.
9. Show that a proposition with n connectives has at most $2n + 1$ subformulas.
10. Show that for PROP we have a unique decomposition theorem: for each non-atomic proposition σ either there are two propositions φ and ψ such that $\sigma = \varphi \Box \psi$, or there is a proposition φ such that $\sigma = \neg\varphi$.
11. (a) Give an inductive definition of the function F, defined by recursion on $PROP$ from the functions H_{at}, $H_{\Box}$, $H_{\neg}$, as a set F^* of pairs.

(b) Formulate and prove for F^* the induction principle.
(c) Prove that F^* is indeed a function on *PROP*.
(d) Prove that it is the unique function on *PROP* satisfying the recursion equations.

2.2 Semantics

The task of interpreting propositional logic is simplified by the fact that the entities considered have a simple structure. The propositions are built up from rough blocks by adding connectives.

The simplest parts (atoms) are of the form "grass is green", "Mary likes Goethe", "$6 - 3 = 2$", which are simply *true* or *false*. We extend this assignment of *truth values* to composite propositions, by reflection on the meaning of the logical connectives.

Let us agree to use 1 and 0 instead of "true" and "false". The problem we are faced with is how to interpret $\varphi \Box \psi$, $\neg\varphi$, given the truth values of φ and ψ.

We will illustrate the solution by considering the in-out table for Messrs. Smith and Jones.

Conjunction A visitor who wants to see both Smith and Jones wants the table to be in the position shown here, i.e.

	in	out
Smith	×	
Jones	×	

"Smith is in" $\wedge$ "Jones is in" is true iff
"Smith is in" is true and "Jones is in" is true.

We write $v(\varphi) = 1$ (resp. 0) for "φ is true" (resp. false). Then the above consideration can be stated as $v(\varphi \wedge \psi) = 1$ iff $v(\varphi) = v(\psi) = 1$, or $v(\varphi \wedge \psi) = \min(v(\varphi), v(\psi))$.

One can also write it in the form of a *truth table*:

$\wedge$	0	1
0	0	0
1	0	1

One reads the truth table as follows: the first argument is taken from the leftmost column and the second argument is taken from the top row.

Disjunction If a visitor wants to see one of the partners, no matter which one, he wants the table to be in one of the positions

	in	out
Smith	×	
Jones		×

	in	out
Smith		×
Jones	×	

	in	out
Smith	×	
Jones	×	

In the last case he can make a choice, but that is no problem; he wants to see at least one of the gentlemen, no matter which one.

In our notation, the interpretation of $\vee$ is given by

$$v(\varphi \vee \psi) = 1 \quad \text{iff} \quad v(\varphi) = 1 \quad \text{or} \quad v(\psi) = 1.$$

Shorter: $v(\varphi \vee \psi) = \max(v(\varphi), v(\psi))$.

In truth table form:

$\vee$	0	1
0	0	1
1	1	1

Negation The visitor who is solely interested in our Smith will state that "Smith is not in" if the table is in the position:

	in	out
Smith		$\times$

So "Smith is not in" is true if "Smith is in" is false. We write this as $v(\neg\varphi) = 1$ iff $v(\varphi) = 0$, or $v(\neg\varphi) = 1 - v(\varphi)$.

In truth table form:

$\neg$	
0	1
1	0

Implication Our legendary visitor has been informed that "Jones is in if Smith is in". Now he can at least predict the following positions of the table:

	in	out
Smith	$\times$	
Jones	$\times$	

	in	out
Smith		$\times$
Jones		$\times$

If the table is in the position

	in	out
Smith	$\times$	
Jones		$\times$

then he knows that the information was false.

The remaining case,

	in	out
Smith		$\times$
Jones	$\times$	

cannot be dealt with in such a simple way. There evidently is no reason to consider the information false, rather "not very helpful", or "irrelevant". However, we have committed ourselves to the position that each statement is true or false, so we decide to call "If Smith is in, then Jones is in" true also in this particular case. The reader should realize that we have made a deliberate choice here; a choice that will prove a happy one in view of the elegance of the system that results. There is no compelling reason, however, to stick to the notion of implication that we just introduced. Various other notions have been studied in the literature; for mathematical purposes our notion (also called "material implication") is, however, perfectly suitable.

Note that there is just one case in which an implication is false (see the truth table below), and one should keep this observation in mind for future application — it helps to cut down calculations.

In our notation the interpretation of implication is given by $v(\varphi \to \psi) = 0$ iff $v(\varphi) = 1$ and $v(\psi) = 0$.

Its truth table is:

$\to$	0	1
0	1	1
1	0	1

Equivalence If our visitor knows that "Smith is in if and only if Jones is in", then he knows that they are either both in, or both out. Hence $v(\varphi \leftrightarrow \psi) = 1$ iff $v(\varphi) = v(\psi)$.

The truth table of $\leftrightarrow$ is:

$\leftrightarrow$	0	1
0	1	0
1	0	1

Falsum An absurdity, such as "$0 \neq 0$", "some odd numbers are even", "I am not myself", cannot be true. So we put $v(\bot) = 0$.

Strictly speaking we should add one more truth table, i.e. the table for $\top$, the opposite of *falsum*.

Verum This symbol stands for a manifestly true proposition such as $1 = 1$; we put $v(\top) = 1$ for all v.

We collect the foregoing in the following definition.

Definition 2.2.1 A mapping $v : PROP \to \{0, 1\}$ is a *valuation* if

$$\begin{aligned}
v(\varphi \wedge \psi) &= \min(v(\varphi), v(\psi)),\\
v(\varphi \vee \psi) &= \max(v(\varphi), v(\psi)),\\
v(\varphi \to \psi) = 0 \quad &\Leftrightarrow \quad v(\varphi) = 1 \text{ and } v(\psi) = 0,\\
v(\varphi \leftrightarrow \psi) = 1 \quad &\Leftrightarrow \quad v(\varphi) = v(\psi),\\
v(\neg\varphi) &= 1 - v(\varphi)\\
v(\bot) &= 0.
\end{aligned}$$

If a valuation is only given for atoms then it is, by virtue of the definition by recursion, possible to extend it to all propositions. Hence we get the following.

Theorem 2.2.2 *If v is a mapping from the atoms into $\{0, 1\}$, satisfying $v(\bot) = 0$, then there exists a unique valuation $[\![\cdot]\!]_v$, such that $[\![\varphi]\!]_v = v(\varphi)$ for atomic φ.*

It has become common practice to denote valuations as defined above by $[\![\varphi]\!]$, so we will adopt this notation. Since $[\![\cdot]\!]$ is completely determined by its values on the atoms, $[\![\varphi]\!]$ is often denoted by $[\![\varphi]\!]_v$. Whenever there is no confusion we will delete the index v.

Theorem 2.2.2 tells us that each of the mappings v and $[\![\cdot]\!]_v$ determines the other one uniquely, therefore we also call v a valuation (or an *atomic valuation*, if necessary). From this theorem it appears that there are many valuations (cf. Exercise 4).

It is also obvious that the *value* $[\![\varphi]\!]_v$ *of* φ *under* v only depends on the values of v on its atomic subformulas.

Lemma 2.2.3 *If* $v(p_i) = v'(p_i)$ *for all* p_i *occurring in* φ, *then* $[\![\varphi]\!]_v = [\![\varphi]\!]_{v'}$.

Proof An easy induction on φ. □

An important subset of *PROP* is that of all propositions φ which are *always true*, i.e. true under all valuations.

Definition 2.2.4

(i) φ is a *tautology* if $[\![\varphi]\!]_v = 1$ for all valuations v.
(ii) $\models \varphi$ stands for "φ is a tautology".
(iii) Let Γ be a set of propositions, then $\Gamma \models \varphi$ iff for all v: ($[\![\psi]\!]_v = 1$ for all $\psi \in \Gamma) \Rightarrow [\![\varphi]\!]_v = 1$.

In words: $\Gamma \models \varphi$ holds iff φ is true under all valuations that make all ψ in Γ true. We say that φ is a semantical consequence of Γ. We write $\Gamma \not\models \varphi$ if $\Gamma \models \varphi$ is not the case.

Convention $\varphi_1, \ldots, \varphi_n \models \psi$ stands for $\{\varphi_1, \ldots, \varphi_n\} \models \psi$.

Note that "$[\![\varphi]\!]_v = 1$ for all v" is another way of saying "$[\![\varphi]\!] = 1$ for all valuations".

Examples

(i) $\models \varphi \to \varphi$; $\quad \models \neg\neg\varphi \to \varphi$; $\quad \models \varphi \vee \psi \leftrightarrow \psi \vee \varphi$,
(ii) $\varphi, \psi \models \varphi \wedge \psi$; $\quad \varphi, \varphi \to \psi \models \psi$; $\quad \varphi \to \psi, \neg\psi \models \neg\varphi$.

One often has to substitute propositions for subformulas; it turns out to be sufficient to define substitution for atoms only.

We write $\varphi[\psi/p_i]$ for the proposition obtained by replacing all occurrences of p_i in φ by ψ. As a matter of fact, substitution of ψ for p_i defines a mapping of *PROP* into *PROP*, which can be given by recursion (on φ).

Definition 2.2.5

$$\varphi[\psi/p_i] = \begin{cases} \varphi & \text{if } \varphi \text{ atomic and } \varphi \neq p_i \\ \psi & \text{if } \varphi = p_i \end{cases}$$

$$(\varphi_1 \Box \varphi_2)[\psi/p_i] = \varphi_1[\psi/p_i] \Box \varphi_2[\psi/p_i]$$
$$(\neg\varphi)[\psi/p_i] = \neg\varphi[\psi/p_i].$$

The following theorem spells out the basic property of the substitution of equivalent propositions.

Theorem 2.2.6 (Substitution Theorem) *If* $\models \varphi_1 \leftrightarrow \varphi_2$, *then* $\models \psi[\varphi_1/p] \leftrightarrow \psi[\varphi_2/p]$, *where* p *is an atom.*

The substitution theorem is actually a consequence of a slightly stronger one.

Lemma 2.2.7 $[\![\varphi_1 \leftrightarrow \varphi_2]\!]_v \leq [\![\psi[\varphi_1/p] \leftrightarrow \psi[\varphi_2/p]]\!]_v$ *and* $\models (\varphi_1 \leftrightarrow \varphi_2) \to (\psi[\varphi_1/p] \leftrightarrow \psi[\varphi_2/p])$.

Proof Induction on ψ. We only have to consider $[\![\varphi_1 \leftrightarrow \varphi_2]\!]_v = 1$ (why?).

- ψ atomic. If $\psi = p$, then $\psi[\varphi_i/p] = \varphi_i$ and the result follows immediately. If $\psi \neq p$, then $\psi[\varphi_i/p] = \psi$, and $[\![\psi[\varphi_1/p] \leftrightarrow \psi[\varphi_2/p]]\!]_v = [\![\psi \leftrightarrow \psi]\!]_v = 1$.
- $\psi = \psi_1 \Box \psi_2$. Induction hypothesis: $[\![\psi_i[\varphi_1/p]]\!]_v = [\![\psi_i[\varphi_2/p]]\!]_v$. Now the value of $[\![(\psi_1 \Box \psi_2)[\varphi_i/p]]\!]_v = [\![\psi_1[\varphi_i/p] \Box \psi_2[\varphi_i/p]]\!]_v$ is uniquely determined by its parts $[\![\psi_j[\varphi_i/p]]\!]_v$, hence $[\![(\psi_1 \Box \psi_2)[\varphi_1/p]]\!]_v = [\![(\psi_1 \Box \psi_2)[\varphi_2/p]]\!]_v$.
- $\psi = \neg\psi_1$. Left to the reader.

The proof of the second part essentially uses the fact that $\models \varphi \to \psi$ iff $[\![\varphi]\!]_v \leq [\![\psi]\!]_v$ for all v (cf. Exercise 6). □

The proof of the substitution theorem now immediately follows. □

The substitution theorem says in plain English that *parts may be replaced by equivalent parts*.

There are various techniques for testing tautologies. One such (rather slow) technique uses truth tables. We give one example:

$$(\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi)$$

φ	ψ	$\neg\varphi$	$\neg\psi$	$\varphi \to \psi$	$\neg\psi \to \neg\varphi$	$(\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi)$
0	0	1	1	1	1	1
0	1	1	0	1	1	1
1	0	0	1	0	0	1
1	1	0	0	1	1	1

The last column consists of 1's only. Since, by Lemma 2.2.3 only the values of φ and ψ are relevant, we had to check 2^2 cases. If there are n (atomic) parts we need 2^n lines.

One can compress the above table a bit, by writing it in the following form:

$(\varphi$	$\rightarrow$	$\psi)$	$\leftrightarrow$	$(\neg\psi$	$\rightarrow$	$\neg\varphi)$
0	1	0	1	1	1	1
0	1	1	1	0	1	1
1	0	0	1	1	0	0
1	1	1	1	0	1	0

Let us make one more remark about the role of the two 0-ary connectives, $\bot$ and $\top$. Clearly, $\models \top \leftrightarrow (\bot \rightarrow \bot)$, so we can define $\top$ from $\bot$. On the other hand, we cannot define $\bot$ from $\top$ and $\rightarrow$; we note that from $\top$ we can never get anything but a proposition equivalent to $\top$ by using $\wedge, \vee, \rightarrow$, but from $\bot$ we can generate $\bot$ and $\top$ by applying $\wedge, \vee, \rightarrow$.

Exercises

1. Check by the truth table method which of the following propositions are tautologies:
   (a) $(\neg\varphi \vee \psi) \leftrightarrow (\psi \rightarrow \varphi)$,
   (b) $\varphi \rightarrow ((\psi \rightarrow \sigma) \rightarrow ((\varphi \rightarrow \psi) \rightarrow (\varphi \rightarrow \sigma)))$,
   (c) $(\varphi \rightarrow \neg\varphi) \leftrightarrow \neg\varphi$,
   (d) $\neg(\varphi \rightarrow \neg\varphi)$,
   (e) $(\varphi \rightarrow (\psi \rightarrow \sigma)) \leftrightarrow ((\varphi \wedge \psi) \rightarrow \sigma)$,
   (f) $\varphi \vee \neg\varphi$ (*principle of the excluded third*),
   (g) $\bot \leftrightarrow (\varphi \wedge \neg\varphi)$,
   (h) $\bot \rightarrow \varphi$ (*ex falso sequitur quodlibet*).
2. Show
   (a) $\varphi \models \varphi$,
   (b) $\varphi \models \psi$ and $\psi \models \sigma \Rightarrow \varphi \models \sigma$,
   (c) $\models \varphi \rightarrow \psi \Leftrightarrow \varphi \models \psi$.
3. Determine $\varphi[\neg p_0 \rightarrow p_3/p_0]$ for $\varphi = p_1 \wedge p_0 \rightarrow (p_0 \rightarrow p_3)$; $\varphi = (p_3 \leftrightarrow p_0) \vee (p_2 \rightarrow \neg p_0)$.
4. Show that there are $2^{\aleph_0}$ valuations.
5. Show

$$
\begin{aligned}
&[\![\varphi \wedge \psi]\!]_v = [\![\varphi]\!]_v \cdot [\![\psi]\!]_v,\\
&[\![\varphi \vee \psi]\!]_v = [\![\varphi]\!]_v + [\![\psi]\!]_v - [\![\varphi]\!]_v \cdot [\![\psi]\!]_v,\\
&[\![\varphi \rightarrow \psi]\!]_v = 1 - [\![\varphi]\!]_v + [\![\varphi]\!]_v \cdot [\![\psi]\!]_v,\\
&[\![\varphi \leftrightarrow \psi]\!]_v = 1 - |[\![\varphi]\!]_v - [\![\psi]\!]_v|.
\end{aligned}
$$

6. Show $[\![\varphi \rightarrow \psi]\!]_v = 1 \Leftrightarrow [\![\varphi]\!]_v \leq [\![\psi]\!]_v$.

2.3 Some Properties of Propositional Logic

On the basis of the previous sections we can already prove a lot of theorems about propositional logic. One of the earliest discoveries in modern propositional logic was its similarity with algebras.

Following Boole, an extensive study of the algebraic properties was made by a number of logicians. The purely algebraic aspects have since then been studied in *Boolean algebra*.

We will just mention a few of those algebraic laws.

Theorem 2.3.1 *The following propositions are tautologies*:

$$(\varphi \vee \psi) \vee \sigma \leftrightarrow \varphi \vee (\psi \vee \sigma) \qquad (\varphi \wedge \psi) \wedge \sigma \leftrightarrow \varphi \wedge (\psi \wedge \sigma)$$

associativity

$$\varphi \vee \psi \leftrightarrow \psi \vee \varphi \qquad \varphi \wedge \psi \leftrightarrow \psi \wedge \varphi$$

commutativity

$$\varphi \vee (\psi \wedge \sigma) \leftrightarrow (\varphi \vee \psi) \wedge (\varphi \vee \sigma) \qquad \varphi \wedge (\psi \vee \sigma) \leftrightarrow (\varphi \wedge \psi) \vee (\varphi \wedge \sigma)$$

distributivity

$$\neg(\varphi \vee \psi) \leftrightarrow \neg\varphi \wedge \neg\psi \qquad \neg(\varphi \wedge \psi) \leftrightarrow \neg\varphi \vee \neg\psi$$

De Morgan's laws

$$\varphi \vee \varphi \leftrightarrow \varphi \qquad \varphi \wedge \varphi \leftrightarrow \varphi$$

idempotency

$$\neg\neg\varphi \leftrightarrow \varphi$$

double negation law

Proof Check the truth tables or do a little computation. For example, De Morgan's law: $[\![\neg(\varphi \vee \psi)]\!] = 1 \Leftrightarrow [\![\varphi \vee \psi]\!] = 0 \Leftrightarrow [\![\varphi]\!] = [\![\psi]\!] = 0 \Leftrightarrow [\![\neg\varphi]\!] = [\![\neg\psi]\!] = 1 \Leftrightarrow [\![\neg\varphi \wedge \neg\psi]\!] = 1$.

So $[\![\neg(\varphi \vee \psi)]\!] = [\![\neg\varphi \wedge \neg\psi]\!]$ for all valuations, i.e. $\models \neg(\varphi \vee \psi) \leftrightarrow \neg\varphi \wedge \neg\psi$.

The remaining tautologies are left to the reader. □

In order to apply the previous theorem in "logical calculations" we need a few more equivalences. This is demonstrated in the simple equivalence $\models \varphi \wedge (\varphi \vee \psi) \leftrightarrow \varphi$ (an exercise for the reader). For, by the distributive law $\models \varphi \wedge (\varphi \vee \psi) \leftrightarrow (\varphi \wedge \varphi) \vee (\varphi \wedge \psi)$ and $\models (\varphi \wedge \varphi) \vee (\varphi \wedge \psi) \leftrightarrow \varphi \vee (\varphi \wedge \psi)$, by idempotency and the substitution theorem. So $\models \varphi \wedge (\varphi \vee \psi) \leftrightarrow \varphi \vee (\varphi \wedge \psi)$. Another application of the distributive law will bring us back to start, so just applying the above laws will not eliminate ψ!

Therefore, we list a few more convenient properties.

Lemma 2.3.2 *If* $\models \varphi \to \psi$, *then*

$$\models \varphi \wedge \psi \leftrightarrow \varphi \quad \textit{and}$$
$$\models \varphi \vee \psi \leftrightarrow \psi$$

Proof By Exercise 6 of Sect. 2.2 $\models \varphi \to \psi$ implies $[\![\varphi]\!]_v \le [\![\psi]\!]_v$ for all v. So $[\![\varphi \wedge \psi]\!]_v = \min([\![\varphi]\!]_v, [\![\psi]\!]_v) = [\![\varphi]\!]_v$ and $[\![\varphi \vee \psi]\!]_v = \max([\![\varphi]\!]_v, [\![\psi]\!]_v) = [\![\psi]\!]_v$ for all v. □

Lemma 2.3.3

(a) $\models \varphi \Rightarrow \models \varphi \wedge \psi \leftrightarrow \psi$,
(b) $\models \varphi \Rightarrow \models \neg\varphi \vee \psi \leftrightarrow \psi$,
(c) $\models \bot \vee \psi \leftrightarrow \psi$,
(d) $\models \top \wedge \psi \leftrightarrow \psi$.

Proof Left to the reader. □

The following theorem establishes some equivalences involving various connectives. It tells us that we can "define" up to logical equivalence all connectives in terms of $\{\vee, \neg\}$, or $\{\rightarrow, \neg\}$, or $\{\wedge, \neg\}$, or $\{\rightarrow, \bot\}$.

That is, we can find e.g. a proposition involving only $\vee$ and $\neg$, which is equivalent to $\varphi \leftrightarrow \psi$, etc.

Theorem 2.3.4

(a) $\models (\varphi \leftrightarrow \psi) \leftrightarrow (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$,
(b) $\models (\varphi \rightarrow \psi) \leftrightarrow (\neg\varphi \vee \psi)$,
(c) $\models \varphi \vee \psi \leftrightarrow (\neg\varphi \rightarrow \psi)$,
(d) $\models \varphi \vee \psi \leftrightarrow \neg(\neg\varphi \wedge \neg\psi)$,
(e) $\models \varphi \wedge \psi \leftrightarrow \neg(\neg\varphi \vee \neg\psi)$,
(f) $\models \neg\varphi \leftrightarrow (\varphi \rightarrow \bot)$,
(g) $\models \bot \leftrightarrow \varphi \wedge \neg\varphi$.

Proof Compute the truth values of the left-hand and right-hand sides. □

We now have enough material to handle logic as if it were algebra. For convenience we write $\varphi \approx \psi$ for $\models \varphi \leftrightarrow \psi$.

Lemma 2.3.5 $\approx$ *is an equivalence relation on PROP*, *i.e.*

$$\begin{array}{l} \varphi \approx \varphi \ (\textit{reflexivity}), \\ \varphi \approx \psi \quad \Rightarrow \quad \psi \approx \varphi \ (\textit{symmetry}), \\ \varphi \approx \psi \textit{ and } \psi \approx \sigma \quad \Rightarrow \quad \varphi \approx \sigma \ (\textit{transitivity}). \end{array}$$

Proof Use $\models \varphi \leftrightarrow \psi$ iff $[\![\varphi]\!]_v = [\![\psi]\!]_v$ for all v. □

We give some examples of algebraic computations, which establish a chain of equivalences.

1. $\models [\varphi \to (\psi \to \sigma)] \leftrightarrow [\varphi \wedge \psi \to \sigma]$,

$$\varphi \to (\psi \to \sigma) \approx \neg\varphi \vee (\psi \to \sigma), \qquad \text{(Theorem 2.3.4(b))}$$
$$\neg\varphi \vee (\psi \to \sigma) \approx \neg\varphi \vee (\neg\psi \vee \sigma), \qquad \text{(Theorem 2.3.4(b) and Subst. Thm.)}$$
$$\neg\varphi \vee (\neg\psi \vee \sigma) \approx (\neg\varphi \vee \neg\psi) \vee \sigma, \qquad \text{(ass.)}$$
$$(\neg\varphi \vee \neg\psi) \vee \sigma \approx \neg(\varphi \wedge \psi) \vee \sigma, \qquad \text{(\textit{De Morgan} and Subst. Thm.)}$$
$$\neg(\varphi \wedge \psi) \vee \sigma \approx (\varphi \wedge \psi) \to \sigma, \qquad \text{(Theorem 2.3.4(b))}$$
$$\text{So } \varphi \to (\psi \to \sigma) \approx (\varphi \wedge \psi) \to \sigma.$$

We now leave out the references to the facts used, and make one long string. We just calculate until we reach a tautology.

2. $\models (\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi)$,
$\neg\psi \to \neg\varphi \approx \neg\neg\psi \vee \neg\varphi \approx \psi \vee \neg\varphi \approx \neg\varphi \vee \psi \approx \varphi \to \psi$
3. $\models \varphi \to (\psi \to \varphi)$,
$\varphi \to (\psi \to \varphi) \approx \neg\varphi \vee (\neg\psi \vee \varphi) \approx (\neg\varphi \vee \varphi) \vee \neg\psi$.

We have seen that $\vee$ and $\wedge$ are associative, therefore we adopt the convention, also used in algebra, to delete brackets in iterated disjunctions and conjunctions; i.e. we write $\varphi_1 \vee \varphi_2 \vee \varphi_3 \vee \varphi_4$, etc. This is all right, since no matter how we restore (syntactically correctly) the brackets, the resulting formula is determined uniquely up to equivalence.

Have we introduced *all* connectives so far? Obviously not. We can easily invent new ones. Here is a famous one, introduced by Sheffer: $\varphi|\psi$ stands for "not both φ and ψ". More precise: $\varphi|\psi$ is given by the following truth table:

Sheffer stroke

|	0	1
0	1	1
1	1	0

Let us say that an n-ary logical connective \$ is *defined* by its truth table, or by its valuation function, if $[\![\$(p_1, \dots, p_n)]\!] = f([\![p_1]\!], \dots, [\![p_n]\!])$ for some function f.

Although we can apparently introduce many new connectives in this way, there are no surprises in stock for us, as all of those connectives are definable in terms of $\vee$ and $\neg$.

Theorem 2.3.6 *For each n-ary connective \$ defined by its valuation function, there is a proposition τ, containing only $p_1, \dots, p_n$, $\vee$ and $\neg$, such that $\models \tau \leftrightarrow \$(p_1, \dots, p_n)$.*

Proof Induction on n. For $n = 1$ there are 4 possible connectives with truth tables

$\$_1$	
0	0
1	0

$\$_2$	
0	1
1	1

$\$_3$	
0	0
1	1

$\$_4$	
0	1
1	0

One easily checks that the propositions $\neg(p \vee \neg p)$, $p \vee \neg p$, p and $\neg p$ will meet the requirements.

Suppose that for all n-ary connectives propositions have been found.
Consider $\$(p_1, \ldots, p_n, p_{n+1})$ with truth table:

p_1	$p_2 \ldots p_n$		p_{n+1}	$\$(p_1, \ldots, p_n, p_{n+1})$
0	0	0	0	i_1
.	.	0	1	i_2
.	0	1	.	.
.	1	1	.	.
0	.	.	.	.
.	1	.	.	.
....	....	....	....	....
1	0	.	.	.
.	.	.	.	.
.	.	.	.	.
.	0	.	.	.
.	1	0	.	.
.	.	0	.	.
1	.	1	0	.
.	.	1	1	$i_{2^{n+1}}$

where $i_k \leq 1$.

We consider two auxiliary connectives $\$_1$ and $\$_2$ defined by

$$\$_1(p_2, \ldots, p_{n+1}) = \$(\bot, p_2, \ldots, p_{n+1}) \quad \text{and}$$
$$\$_2(p_2, \ldots, p_{n+1}) = \$(\top, p_2, \ldots, p_{n+1}), \quad \text{where } \top = \neg\bot$$

(as given by the upper and lower halves of the above table).

By the induction hypothesis there are propositions σ_1 and σ_2, containing only $p_2, \ldots, p_{n+1}, \vee$ and $\neg$ so that $\models \$_i(p_2, \ldots, p_{n+1}) \leftrightarrow \sigma_i$.

From those two propositions we can construct the proposition τ:

$$[\tau := (p_1 \to \sigma_2) \wedge (\neg p_1 \to \sigma_1).$$

Claim $\models \$(p_1, \ldots, p_{n+1}) \leftrightarrow \tau$.

If $[\![p_1]\!]_v = 0$, then $[\![p_1 \to \sigma_2]\!]_v = 1$, so $[\![\tau]\!]_v = [\![\neg p_1 \to \sigma_1]\!]_v = [\![\sigma_1]\!]_v = [\![\$_1(p_2, \ldots, p_{n+1})]\!]_v = [\![\$(p_1, p_2, \ldots, p_{n+1})]\!]_v$, using $[\![p_1]\!]_v = 0 = [\![\bot]\!]_v$.

The case $[\![p_1]\!]_v = 1$ is similar.

Now expressing $\to$ and $\wedge$ in terms of $\vee$ and $\neg$ (2.3.4), we have $[\![\tau']\!] = [\![\$(p_1, \ldots, p_{n+1})]\!]$ for all valuations (another use of Lemma 2.3.5), where $\tau' \approx \tau$ and τ' contains only the connectives $\vee$ and $\neg$. □

For another solution see Exercise 7.

The above theorem and Theorem 2.3.4 are pragmatic justifications for our choice of the truth table for $\to$: we get an extremely elegant and useful theory. Theorem 2.3.6 is usually expressed by saying that $\vee$ and $\neg$ form a *functionally complete* set of connectives. Likewise $\wedge, \neg$ and $\rightarrow, \neg$ and $\bot, \rightarrow$ form functionally complete sets.

In analogy to the $\sum$ and $\prod$ from algebra we introduce finite disjunctions and conjunctions.

Definition 2.3.7

$$\begin{cases} \bigwedge_{i\le 0}\varphi_i = \varphi_0 \\ \bigwedge_{i\le n+1}\varphi_i = \bigwedge_{i\le n}\varphi_i \wedge \varphi_{n+1} \end{cases} \qquad \begin{cases} \bigvee_{i\le 0}\varphi_i = \varphi_0 \\ \bigvee_{i\le n+1}\varphi_i = \bigvee_{i\le n}\varphi_i \vee \varphi_{n+1} \end{cases}$$

Definition 2.3.8 If $\varphi = \bigwedge_{i\le n}\bigvee_{j\le m_i}\varphi_{ij}$, where φ_{ij} is atomic or the negation of an atom, then φ is a *conjunctive normal form*. If $\varphi = \bigvee_{i\le n}\bigwedge_{j\le m_i}\varphi_{ij}$, where φ_{ij} is atomic or the negation of an atom, then φ is a *disjunctive normal form*.

The normal forms are analogous to the well-known normal forms in algebra: ax^2+byx is “normal”, whereas $x(ax+by)$ is not. One can obtain normal forms by simply “multiplying”, i.e. repeated application of distributive laws. In algebra there is only one “normal form”; in logic there is a certain duality between $\wedge$ and $\vee$, so that we have two normal form theorems.

Theorem 2.3.9 *For each φ there are conjunctive normal forms $\varphi^\wedge$ and disjunctive normal forms $\varphi^\vee$, such that $\models \varphi \leftrightarrow \varphi^\wedge$ and $\models \varphi \leftrightarrow \varphi^\vee$.*

Proof First eliminate all connectives other than $\bot, \wedge, \vee$ and $\neg$. Then prove the theorem by induction on the resulting proposition in the restricted language of $\bot, \wedge, \vee$ and $\neg$. In fact, $\bot$ plays no role in this setting; it could just as well be ignored.

(a) φ is atomic. Then $\varphi^\wedge = \varphi^\vee = \varphi$.
(b) $\varphi = \psi \wedge \sigma$. Then $\varphi^\wedge = \psi^\wedge \wedge \sigma^\wedge$. In order to obtain a disjunctive normal form we consider $\psi^\vee = \bigvee \psi_i$, $\sigma^\vee = \bigvee \sigma_j$, where the ψ_i’s and σ_j’s are conjunctions of atoms and negations of atoms.

Now $\varphi = \psi \wedge \sigma \approx \psi^\vee \wedge \sigma^\vee \approx \bigvee_{i,j}(\psi_i \wedge \sigma_j)$.
The last proposition is in normal form, so we equate $\varphi^\vee$ to it.

(c) $\varphi = \psi \vee \sigma$. Similar to (b).
(d) $\varphi = \neg\psi$. By the induction hypothesis ψ has normal forms $\psi^\vee$ and $\psi^\wedge$. $\neg\psi \approx \neg\psi^\wedge \approx \neg\bigvee\bigwedge\psi_{ij} \approx \bigwedge\bigvee\neg\psi_{ij} \approx \bigwedge\bigvee\psi'_{ij}$, where $\psi'_{ij} = \neg\psi_{ij}$ if ψ_{ij} is atomic, and $\psi_{ij} = \neg\psi'_{ij}$ if ψ_{ij} is the negation of an atom. (Observe $\neg\neg\psi_{ij} \approx \psi_{ij}$.) Clearly $\bigwedge\bigvee\psi'_{ij}$ is a conjunctive normal form for φ. The disjunctive normal form is left to the reader.

For another proof of the normal form theorems see Exercise 7. □

When looking at the algebra of logic in Theorem 2.3.1, we saw that $\vee$ and $\wedge$ behaved in a very similar way, to the extent that the same laws hold for both. We will make this “duality” precise. For this purpose we consider a language with only the connectives $\vee, \wedge$ and $\neg$.

Definition 2.3.10 Define an auxiliary mapping * : *PROP* → *PROP* recursively by

$$\begin{aligned}\varphi^* &= \neg\varphi \quad \text{if } \varphi \text{ is atomic,}\\ (\varphi \wedge \psi)^* &= \varphi^* \vee \psi^*,\\ (\varphi \vee \psi)^* &= \varphi^* \wedge \psi^*,\\ (\neg\varphi)^* &= \neg\varphi^*.\end{aligned}$$

Example $((p_0 \wedge \neg p_1) \vee p_2)^* = (p_0 \wedge \neg p_1)^* \wedge p_2^* = (p_0^* \vee (\neg p_1)^*) \wedge \neg p_2 = (\neg p_0 \vee \neg p_1^*) \wedge \neg p_2 = (\neg p_0 \vee \neg\neg p_1) \wedge \neg p_2 \approx (\neg p_0 \vee p_1) \wedge \neg p_2$.

Note that the effect of the *-translation boils down to taking the negation and applying De Morgan's laws.

Lemma 2.3.11 $[\![\varphi^*]\!] = [\![\neg\varphi]\!]$.

Proof Induction on φ. For atomic φ $[\![\varphi^*]\!] = [\![\neg\varphi]\!]$. $[\![(\varphi \wedge \psi)^*]\!] = [\![\varphi^* \vee \psi^*]\!] = [\![\neg\varphi \vee \neg\psi]\!]) = [\![\neg(\varphi \wedge \psi)]\!])$. $[\![(\varphi \vee \psi)^*]\!]$ and $[\![(\neg\varphi)^*]\!]$ are left to the reader. □

Corollary 2.3.12 $\models \varphi^* \leftrightarrow \neg\varphi$.

Proof The proof is immediate from Lemma 2.3.11. □

So far this is not the proper duality we have been looking for. We really just want to interchange $\wedge$ and $\vee$. So we introduce a new translation.

Definition 2.3.13 The duality mapping ^{d} : *PROP* → *PROP* is recursively defined by

$$\begin{aligned}\varphi^d &= \varphi \quad \text{for } \varphi \text{ atomic,}\\ (\varphi \wedge \psi)^d &= \varphi^d \vee \psi^d,\\ (\varphi \vee \psi)^d &= \varphi^d \wedge \psi^d,\\ (\neg\varphi)^d &= \neg\varphi^d.\end{aligned}$$

Theorem 2.3.14 (Duality Theorem) $\models \varphi \leftrightarrow \psi \;\Leftrightarrow\; \models \varphi^d \leftrightarrow \psi^d$.

Proof We use the $*$-translation as an intermediate step. Let us introduce the notion of simultaneous substitution to simplify the proof.

$\sigma[\tau_0, \ldots, \tau_n / p_0, \ldots, p_n]$ is obtained by substituting τ_i for p_i for all $i \leq n$ simultaneously (see Exercise 15). Observe that $\varphi^* = \varphi^d[\neg p_0, \ldots, \neg p_n / p_0, \ldots, p_n]$, so $\varphi^*[\neg p_0, \ldots, \neg p_n / p_0, \ldots, p_n] = \varphi^d[\neg\neg p_0, \ldots, \neg\neg p_n / p_0, \ldots, p_n]$, where the atoms of φ occur among the $p_0, \ldots, p_n$.

By the Substitution Theorem $\models \varphi^d \leftrightarrow \varphi^*[\neg p_0, \ldots, \neg p_n / p_0, \ldots, p_n]$. The same equivalence holds for ψ.

By Corollary 2.3.12 $\models \varphi^* \leftrightarrow \neg\varphi$, $\models \psi^* \leftrightarrow \neg\psi$. Since $\models \varphi \leftrightarrow \psi$, also $\models \neg\varphi \leftrightarrow \neg\psi$. Hence $\models \varphi^* \leftrightarrow \psi^*$, and therefore $\models \varphi^*[\neg p_0, \dots, \neg p_n/p_0, \dots, p_n] \leftrightarrow \psi^*[\neg p_0, \dots, \neg p_n/p_0, \dots, p_n]$.

Using the above relation between φ^d and φ^* we now obtain $\models \varphi^d \leftrightarrow \psi^d$. The converse follows immediately, as $\varphi^{dd} = \varphi$. □

The Duality Theorem gives us one identity for free for each identity we establish.

Exercises

1. Show by "algebraic" means:

$$
\begin{array}{ll}
\models & (\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi), \quad \text{contraposition,} \\
\models & (\varphi \to \psi) \wedge (\psi \to \sigma) \to (\varphi \to \sigma), \quad \text{transitivity of } \to, \\
\models & (\varphi \to (\psi \wedge \neg\psi)) \to \neg\varphi, \\
\models & (\varphi \to \neg\varphi) \to \neg\varphi, \\
\models & \neg(\varphi \wedge \neg\varphi), \\
\models & \varphi \to (\psi \to \varphi \wedge \psi), \\
\models & ((\varphi \to \psi) \to \varphi) \to \varphi, \quad \text{Peirce's law.}
\end{array}
$$

2. Simplify the following propositions (i.e. find a simpler equivalent proposition):

   (a) $(\varphi \to \psi) \wedge \varphi$, (b) $(\varphi \to \psi) \vee \neg\varphi$, (c) $(\varphi \to \psi) \to \psi$,
   (d) $\varphi \to (\varphi \wedge \psi)$, (e) $(\varphi \wedge \psi) \vee \varphi$, (f) $(\varphi \to \psi) \to \varphi$.

3. Show that $\{\neg\}$ is not a functionally complete set of connectives. Idem for $\{\to, \vee\}$ (hint: show that for each formula φ with only $\to$ and $\vee$ there is a valuation v such that $[\![\varphi]\!]_v = 1$).
4. Show that the Sheffer stroke, $|$, forms a functionally complete set (hint: $\models \neg\varphi \leftrightarrow \varphi \mid \varphi$).
5. Show that the connective $\downarrow$ (φ *nor* ψ), with valuation function $[\![\varphi\downarrow\psi]\!] = 1$ iff $[\![\varphi]\!] = [\![\psi]\!] = 0$, forms a functionally complete set.
6. Show that $|$ and $\downarrow$ are the only binary connectives \$ such that {\$} is functionally complete.
7. The functional completeness of $\{\vee, \neg\}$ can be shown in an alternative way. Let \$ be an n-ary connective with valuation function $[\![\$(p_1, \dots, p_n)]\!] = f([\![p_1]\!], \dots, [\![p_n]\!])$. We want a proposition τ (in $\vee, \neg$) such that $[\![\tau]\!] = f([\![p_1]\!], \dots, [\![p_n]\!])$.

   Suppose $f([\![p_1]\!], \dots, [\![p_n]\!]) = 1$ at least once. Consider all tuples $([\![p_1]\!], \dots, [\![p_n]\!])$ with $f([\![p_1]\!], \dots, [\![p_n]\!]) = 1$ and form corresponding conjunctions $\bar{p}_1 \wedge \bar{p}_2 \wedge \dots \wedge \bar{p}_n$ such that $\bar{p}_i = p_i$ if $[\![p_i]\!] = 1$, $\bar{p}_i = \neg p_i$ if $[\![p_i]\!] = 0$. Then show $\models (\bar{p}_1^1 \wedge \bar{p}_2^1 \wedge \dots \wedge \bar{p}_n^1) \vee \dots \vee (\bar{p}_1^k \wedge \bar{p}_2^k \wedge \dots \wedge \bar{p}_n^k) \leftrightarrow \$(p_1, \dots, p_n)$, where the disjunction is taken over all n-tuples such that $f([\![p_1]\!], \dots, [\![p_n]\!]) = 1$.

   Alternatively, we can consider the tuples for which $f([\![p_1]\!], \dots, [\![p_n]\!]) = 0$. Carry out the details. Note that this proof of the functional completeness at the same time proves the normal form theorems.

8. Let the ternary connective \$ be defined by $[\![\$(\varphi_1,\varphi_2,\varphi_3)]\!]=1 \Leftrightarrow [\![\varphi_1]\!]+[\![\varphi_2]\!]+[\![\varphi_3]\!]\geq 2$ (the majority connective). Express \$ in terms of $\vee$ and $\neg$.
9. Let the binary connective # be defined by

#	0	1
0	0	1
1	1	0

   Express # in terms of $\vee$ and $\neg$.
10. Determine conjunctive and disjunctive normal forms for $\neg(\varphi \leftrightarrow \psi)$, $((\varphi\to\psi)\to\psi)\to\psi$, $(\varphi\to(\varphi\wedge\neg\psi))\wedge(\psi\to(\psi\wedge\neg\varphi))$.
11. Give a criterion for a conjunctive normal form to be a tautology.
12. Prove

$$\bigwedge_{i\leq n}\varphi_i \vee \bigwedge_{j\leq m}\psi_j \approx \bigwedge_{\substack{i\leq n\\ j\leq m}}(\varphi_i\vee\psi_j)$$

   and

$$\bigvee_{i\leq n}\varphi_i \wedge \bigvee_{j\leq m}\psi_j \approx \bigvee_{\substack{i\leq n\\ j\leq m}}(\varphi_i\wedge\psi_j).$$

13. The set of all valuations, thought of as the set of all 0–1-sequences, forms a topological space, called the Cantor space $\mathcal{C}$. The basic open sets are finite unions of sets of the form $\{v \mid [\![p_{i_1}]\!]_v=\cdots=[\![p_{i_n}]\!]_v=1 \text{ and } [\![p_{j_1}]\!]_v=\cdots=[\![p_{j_m}]\!]_v=0\}$, $i_k\neq j_p$ for $k\leq n$; $p\leq m$.

   Define a function $[\![\ ]\!]: PROP\to\mathcal{P}(\mathcal{C})$ (subsets of the Cantor space) by: $[\![\varphi]\!]=\{v\mid [\![\varphi]\!]_v=1\}$.

   (a) Show that $[\![\varphi]\!]$ is a basic open set (which is also closed),
   (b) $[\![\varphi\vee\psi]\!]=[\![\varphi]\!]\cup[\![\psi]\!]$; $[\![\varphi\wedge\psi]\!]=[\![\varphi]\!]\cap[\![\psi]\!]$; $[\![\neg\varphi]\!]=[\![\varphi]\!]^c$,
   (c) $\models\varphi\Leftrightarrow[\![\varphi]\!]=C$; $[\![\bot]\!]=\emptyset$; $\models\varphi\to\psi\Leftrightarrow[\![\varphi]\!]\subseteq[\![\psi]\!]$.
   Extend the mapping to sets of propositions Γ by $[\![\Gamma]\!]=\{v\mid [\![\varphi]\!]_v=1$ for all $\varphi\in\Gamma\}$. Note that $[\![\Gamma]\!]$ is closed.
   (d) $\Gamma\models\varphi\Leftrightarrow[\![\Gamma]\!]\subseteq[\![\varphi]\!]$.
14. We can view the relation $\models\varphi\to\psi$ as a kind of ordering. Put $\varphi\sqsubset\psi :=$ $\models\varphi\to\psi$ and $\not\models\psi\to\varphi$.
   (i) For each φ,ψ such that $\varphi\sqsubset\psi$, find σ with $\varphi\sqsubset\sigma\sqsubset\psi$.
   (ii) Find $\varphi_1,\varphi_2,\varphi_3,\ldots$ such that $\varphi_1\sqsubset\varphi_2\sqsubset\varphi_3\sqsubset\varphi_4\sqsubset\ldots$,
   (iii) and show that for each φ,ψ with φ and ψ incomparable, there is a least σ with $\varphi,\psi\sqsubset\sigma$.
15. Give a recursive definition of the simultaneous substitution $\varphi[\psi,\ldots,\psi_n/p_1,\ldots,p_n]$ and formulate and prove the appropriate analogue of the Substitution Theorem (Theorem 2.2.6).

2.4 Natural Deduction

In the preceding sections we have adopted the view that propositional logic is based on truth tables; i.e. we have looked at logic from a semantical point of view. This, however, is not the only possible point of view. If one thinks of logic as a codification of (exact) reasoning, then it should stay close to the practice of inference making, instead of basing itself on the notion of truth. We will now explore the non-semantic approach, by setting up a system for deriving conclusions from premises. Although this approach is of a formal nature, i.e. it abstains from interpreting the statements and rules, it is advisable to keep some interpretation in mind. We are going to introduce a number of derivation rules, which are, in a way, the atomic steps in a derivation. These derivation rules are designed (by Gentzen), to render the intuitive meaning of the connectives as faithfully as possible.

There is one minor problem, which at the same time is a major advantage, namely: our rules express the constructive meaning of the connectives. This advantage will not be exploited now, but it is good to keep it in mind when dealing with logic (it is exploited in intuitionistic logic).

One small example: the principle of the excluded third tells us that $\models \varphi \vee \neg\varphi$, i.e., assuming that φ is a definite mathematical statement, either it or its negation must be true. Now consider some unsolved problem, e.g. Riemann's hypothesis, call it R. Then either R is true, or $\neg R$ is true. However, we do not know which of the two is true, so the constructive content of $R \vee \neg R$ is nil. Constructively, one would require a method to find out which of the alternatives holds.

The propositional connective which has a strikingly different meaning in a constructive and in a non-constructive approach is the disjunction. Therefore we restrict our language for the moment to the connectives $\wedge$, $\to$ and $\bot$. This is no real restriction as $\{\to, \bot\}$ is a functionally complete set.

Our derivations consist of very simple steps, such as "from φ and $\varphi \to \psi$ conclude ψ", written as:

$$\frac{\varphi \quad \varphi \to \psi}{\psi}$$

The propositions above the line are *premises*, and the one below the line is the *conclusion*. The above example *eliminated* the connective $\to$. We can also *introduce* connectives. The derivation rules for $\wedge$ and $\to$ are separated into

Introduction Rules	Elimination Rules
$(\wedge I)\quad \dfrac{\varphi \quad \psi}{\varphi \wedge \psi}\wedge I$	$(\wedge E)\quad \dfrac{\varphi \wedge \psi}{\varphi}\wedge E \qquad \dfrac{\varphi \wedge \psi}{\psi}\wedge E$
$(\to I)\quad \begin{array}{c}[\varphi] \\ \vdots \\ \dfrac{\psi}{\varphi \to \psi}\to I\end{array}$	$(\to E)\quad \dfrac{\varphi \quad \varphi \to \psi}{\psi}\to E$

We have two rules for $\perp$, both of which eliminate $\perp$, but introduce a formula.

$$(\perp)\ \frac{\perp}{\varphi}\perp \qquad (\text{RAA})\ \begin{array}{c}[\neg\varphi]\\ \vdots\\ \dfrac{\perp}{\varphi}\,\text{RAA}\end{array}$$

As usual "$\neg\varphi$" is used here as an abbreviation for "$\varphi \to \perp$".

The rules for $\wedge$ are evident: if we have φ and ψ we may conclude $\varphi \wedge \psi$, and if we have $\varphi \wedge \psi$ we may conclude φ (or ψ). The introduction rule for implication has a different form. It states that, if we can derive ψ from φ (as a hypothesis), then we may conclude $\varphi \to \psi$ (without the hypothesis φ). This agrees with the intuitive meaning of implication: $\varphi \to \psi$ means "ψ follows from φ". We have written the rule ($\to$ I) in the above form to suggest a derivation. The notation will become clearer after we have defined derivations. For the time being we will write the premises of a rule in the order that suits us best, later we will become more fastidious.

The rule ($\to E$) is also evident on the meaning of implication. If φ is given and we know that ψ follows from φ, then we also have ψ. The *falsum rule*, ($\perp$), expresses that from an absurdity we can derive everything (ex falso sequitur quodlibet), and the *reductio ad absurdum rule*, (RAA), is a formulation of the *principle of proof by contradiction*: if one derives a contradiction from the hypothesis $\neg\varphi$, then one has a derivation of φ (without the hypothesis $\neg\varphi$, of course). In both ($\to I$) and (RAA) hypotheses disappear, which is indicated by the striking out of the hypothesis. We say that such a hypothesis is *canceled*. Let us digress for a moment on the cancellation of hypotheses. We first consider implication introduction. There is a well-known theorem in plane geometry which states that "if a triangle is isosceles, then the angles opposite the equal sides are equal to one another" (Euclid's Elements, Book I, Proposition 5). This is shown as follows: we suppose that we have an isosceles triangle and then, in a number of steps, we deduce that the angles at the base are equal. Thence we conclude that *the angles at the base are equal if the triangle is isosceles*.

Query 1: do we still need the hypothesis that the triangle is isosceles? Of course not! We have, so to speak, incorporated this condition in the statement itself. It is precisely the role of conditional statements, such as "if it rains I will use my umbrella", to get rid of the obligation to require (or verify) the condition. In abstracto: if we can deduce ψ using the hypothesis φ, then $\varphi \to \psi$ is the case *without the hypothesis* φ (there may be other hypotheses, of course).

Query 2: is it forbidden to maintain the hypothesis? Answer: no, but it clearly is superfluous. As a matter of fact we usually experience superfluous conditions as confusing or even misleading, but that is rather a matter of the psychology of problem solving than of formal logic. Usually we want the best possible result, and it is intuitively clear that the more hypotheses we state for a theorem, the weaker our result is. Therefore we will as a rule cancel as many hypotheses as possible.

In the case of (RAA) we also deal with cancellation of hypotheses. Again, let us consider an example.

In analysis we introduce the notion of a *convergent sequence* (a_n) and subsequently the notion "a is a limit of (a_n)". The next step is to prove that for each convergent sequence there is a unique limit; we are interested in the part of the proof that shows that there is at most one limit. Such a proof may run as follows: we suppose that there are two distinct limits a and a', and from this hypothesis, $a \neq a'$, we derive a contradiction. Conclusion: $a = a'$. In this case we of course drop the hypothesis $a \neq a'$; this time it is not a case of being superfluous, but of being in conflict! So, both in the case $(\to I)$ and in (RAA), it is sound practice to cancel all occurrences of the hypothesis concerned.

In order to master the technique of natural deduction, and to become familiar with the technique of cancellation, one cannot do better than to look at a few concrete cases. So before we go on to the notion of *derivation* we consider a few examples.

$$\mathbf{I}\qquad \dfrac{\dfrac{\dfrac{[\varphi\wedge\psi]^1}{\psi}\wedge E \qquad \dfrac{[\varphi\wedge\psi]^1}{\varphi}\wedge E}{\psi\wedge\varphi}\wedge I}{\varphi\wedge\psi\to\psi\wedge\varphi}\to I_1 \qquad\qquad \mathbf{II}\qquad \dfrac{\dfrac{\dfrac{[\varphi]^2\quad[\varphi\to\bot]^1}{\bot}\to E}{(\varphi\to\bot)\to\bot}\to I_1}{\varphi\to((\varphi\to\bot)\to\bot)}\to I_2$$

$$\mathbf{III}\qquad \dfrac{\dfrac{\dfrac{\dfrac{[\varphi\wedge\psi]^1}{\psi}\wedge E \qquad \dfrac{\dfrac{[\varphi\wedge\psi]^1}{\varphi}\wedge E \qquad [\varphi\to(\psi\to\sigma)]^2}{\psi\to\sigma}\to E}{\sigma}\to E}{\varphi\wedge\psi\to\sigma}\to I_1}{(\varphi\to(\psi\to\sigma))\to(\varphi\wedge\psi\to\sigma)}\to I_2$$

If we use the customary abbreviation "$\neg\varphi$" for "$\varphi\to\bot$", we can bring some derivations into a more convenient form. (Recall that $\neg\varphi$ and $\varphi\to\bot$, as given in 2.2, are semantically equivalent.) We rewrite derivation II using the abbreviation:

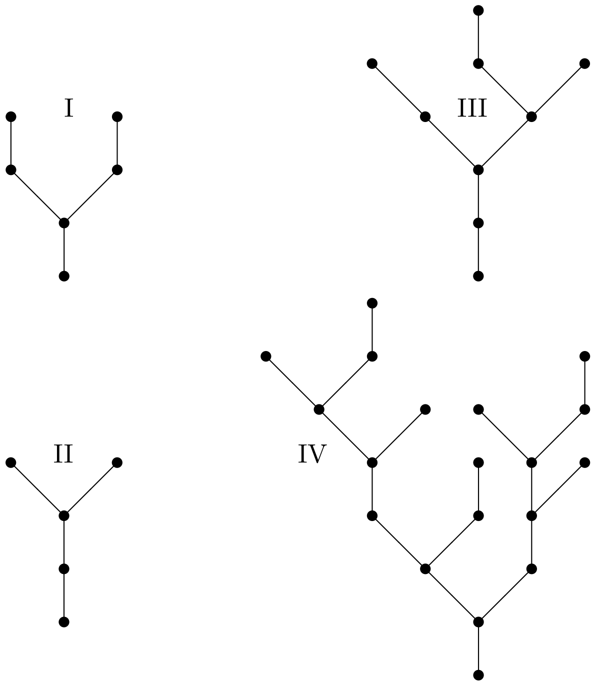
$$\mathbf{II}'\qquad \dfrac{\dfrac{\dfrac{[\varphi]^2\quad[\neg\varphi]^1}{\bot}\to E}{\neg\neg\varphi}\to I_1}{\varphi\to\neg\neg\varphi}\to I_2$$

In the following example we use the negation sign and also the bi-implication; $\varphi\leftrightarrow\psi$ for $(\varphi\to\psi)\wedge(\psi\to\varphi)$.

IV

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^1 \quad \dfrac{[\varphi\leftrightarrow\neg\varphi]^3}{\varphi\to\neg\varphi}\wedge E}{\neg\varphi}\to E \qquad [\varphi]^1}{\bot}\to E}{\neg\varphi}\to I_1 \qquad \dfrac{[\varphi\leftrightarrow\neg\varphi]^3}{\neg\varphi\to\varphi}\wedge E}{\varphi}\to E \qquad \dfrac{\dfrac{\dfrac{[\varphi]^2 \quad \dfrac{[\varphi\leftrightarrow\neg\varphi]^3}{\varphi\to\neg\varphi}\wedge E}{\neg\varphi}\to E \qquad [\varphi]^2}{\bot}\to E}{\neg\varphi}\to I_2}{\dfrac{\bot}{\neg(\varphi\leftrightarrow\neg\varphi)}\to I_3}\to E$$

The examples show us that derivations have the form of trees. We show the trees below:



One can just as well present derivations as (linear) strings of propositions. We will stick, however, to the tree form, the idea being that what comes naturally in tree form should not be put in a linear straightjacket.

We now strive to define the notion of *derivation* in general. We will use an inductive definition to produce trees.

Notation If $\begin{array}{c}\mathcal{D}\\ \varphi\end{array}, \begin{array}{c}\mathcal{D}'\\ \varphi'\end{array}$ are derivations with conclusions φ, φ', then $\begin{array}{c}\mathcal{D}\\ \dfrac{\varphi}{\psi}\end{array}, \begin{array}{c}\mathcal{D}\;\mathcal{D}'\\ \dfrac{\varphi\;\varphi'}{\psi}\end{array}$ are derivations obtained by applying a derivation rule to φ (and φ and φ'). The cancellation of a hypothesis is indicated as follows: if $\begin{array}{c}\psi\\ \mathcal{D}\\ \varphi\end{array}$ is a derivation with hypothesis ψ, then $\begin{array}{c}[\psi]\\ \mathcal{D}\\ \varphi\\ \hline \sigma\end{array}$ is a derivation with ψ canceled.

With respect to the cancellation of hypotheses, we note that one does not necessarily cancel *all* occurrences of such a proposition ψ. This clearly is justified, as one feels that adding hypotheses does not make a proposition underivable (irrelevant information may always be added). It is a matter of prudence, however, to cancel as much as possible. Why carry more hypotheses than necessary?

Furthermore one may apply ($\rightarrow$ I) if there is no hypothesis available for cancellation; e.g. $\frac{\varphi}{\psi \rightarrow \varphi} \rightarrow I$ is a correct derivation, using just $(\rightarrow I)$. To sum up: given a derivation tree of ψ (or $\bot$), we obtain a derivation tree of $\varphi \rightarrow \psi$ (or φ) at the bottom of the tree and strike out some (or all) occurrences, if any, of φ (or $\neg\varphi$) on top of a tree.

A few words on the practical use of natural deduction: if you want to give a derivation for a proposition it is advisable to devise some kind of strategy, just as in a game. Suppose that you want to show $[\varphi \wedge \psi \rightarrow \sigma] \rightarrow [\varphi \rightarrow (\psi \rightarrow \sigma)]$ (Example III), then (since the proposition is an implicational formula) the rule $(\rightarrow I)$ suggests itself. So try to derive $\varphi \rightarrow (\psi \rightarrow \sigma)$ from $\varphi \wedge \psi \rightarrow \sigma$.

Now we know where to start and where to go to. To make use of $\varphi \wedge \psi \rightarrow \sigma$ we want $\varphi \wedge \psi$ (for $(\rightarrow E)$), and to get $\varphi \rightarrow (\psi \rightarrow \sigma)$ we want to derive $\psi \rightarrow \sigma$ from φ. So we may add φ as a hypothesis and look for a derivation of $\psi \rightarrow \sigma$. Again, this asks for a derivation of σ from ψ, so add ψ as a hypothesis and look for a derivation of σ. By now we have the following hypotheses available: $\varphi \wedge \psi \rightarrow \sigma$, φ and ψ. Keeping in mind that we want to eliminate $\varphi \wedge \psi$ it is evident what we should do. The derivation III shows in detail how to carry out the derivation. After making a number of derivations one gets the practical conviction that one should first take propositions apart from the bottom upwards, and then construct the required propositions by putting together the parts in a suitable way. This practical conviction is confirmed by the *Normalization Theorem*, to which we will return later. There is a particular point which tends to confuse novices:

$$\begin{array}{c}[\varphi]\\ \vdots\\ \bot\\ \hline \neg\varphi\end{array} \rightarrow I \qquad \text{and} \qquad \begin{array}{c}[\neg\varphi]\\ \vdots\\ \bot\\ \hline \varphi\end{array} \text{RAA}$$

look very much alike. Are they not both cases of reductio ad absurdum? As a matter of fact the leftmost derivation tells us (informally) that the assumption of φ leads to a contradiction, so φ *cannot be the case*. This is in our terminology the meaning of "not φ". The rightmost derivation tells us that the assumption of $\neg\varphi$ leads to a contradiction, hence (by the same reasoning) $\neg\varphi$ cannot be the case. So, on account of the meaning of negation, we only would get $\neg\neg\varphi$. It is by no means clear that $\neg\neg\varphi$ is equivalent to φ (indeed, this is denied by the intuitionists), so it is an extra property of our logic. (This is confirmed in a technical sense: $\neg\neg\varphi \to \varphi$ is not derivable in the system without RAA.)

We now return to our theoretical notions.

Definition 2.4.1 The set of derivations is the smallest set X such that

(1) The one-element tree φ belongs to X for all $\varphi \in PROP$.

(2$\wedge$) If $\begin{array}{c}\mathcal{D}\\ \varphi\end{array}, \begin{array}{c}\mathcal{D}'\\ \varphi'\end{array} \in X$, then $\dfrac{\begin{array}{cc}\mathcal{D} & \mathcal{D}'\\ \varphi & \varphi'\end{array}}{\varphi\wedge\varphi'} \in X$.

If $\begin{array}{c}\mathcal{D}\\ \varphi\wedge\psi\end{array} \in X$, then $\dfrac{\begin{array}{c}\mathcal{D}\\ \varphi\wedge\psi\end{array}}{\varphi}, \dfrac{\begin{array}{c}\mathcal{D}\\ \varphi\wedge\psi\end{array}}{\psi} \in X$.

(2$\to$) If $\begin{array}{c}\varphi\\ \mathcal{D}\\ \psi\end{array} \in X$, then $\dfrac{\begin{array}{c}[\varphi]\\ \mathcal{D}\\ \psi\end{array}}{\varphi\to\psi} \in X$.

If $\begin{array}{c}\mathcal{D}\\ \varphi\end{array}, \begin{array}{c}\mathcal{D}'\\ \varphi\to\psi\end{array} \in X$, then $\dfrac{\begin{array}{cc}\mathcal{D} & \mathcal{D}'\\ \varphi & \varphi\to\psi\end{array}}{\psi} \in X$.

(2$\bot$) If $\begin{array}{c}\mathcal{D}\\ \bot\end{array} \in X$, then $\dfrac{\begin{array}{c}\mathcal{D}\\ \bot\end{array}}{\varphi} \in X$.

If $\begin{array}{c}\neg\varphi\\ \mathcal{D}\\ \bot\end{array} \in X$, then $\dfrac{\begin{array}{c}[\neg\varphi]\\ \mathcal{D}\\ \bot\end{array}}{\varphi} \in X$.

The bottom formula of a derivation is called its *conclusion*. Since the class of derivations is inductively defined, we can mimic the results of Sect. 2.1.

For example, we have a *principle of induction on* $\mathcal{D}$: let A be a property. If $A(\mathcal{D})$ holds for one-element derivations and A is preserved under the clauses (2$\wedge$), (2 $\to$) and (2 $\bot$), then $A(\mathcal{D})$ holds for all derivations. Likewise we can define mappings on the set of derivations by recursion (cf. Exercises 6, 7, 9).

Definition 2.4.2 The relation $\Gamma \vdash \varphi$ between sets of propositions and propositions is defined as follows: there is a derivation with conclusion φ and with all (uncanceled) hypotheses in Γ. (See also Exercise 6.)

We say that φ is *derivable* from Γ. Note that by definition Γ may contain many superfluous "hypotheses". The symbol $\vdash$ is called the *turnstile*.

If $\Gamma = \emptyset$, we write $\vdash \varphi$, and we say that φ is a theorem.

We could have avoided the notion of "derivation" and taken instead the notion of "derivability" as fundamental, see Exercise 10. The two notions, however, are closely related.

Lemma 2.4.3

(a) $\Gamma \vdash \varphi$ *if* $\varphi \in \Gamma$,
(b) $\Gamma \vdash \varphi, \Gamma' \vdash \psi \Rightarrow \Gamma \cup \Gamma' \vdash \varphi \wedge \psi$,
(c) $\Gamma \vdash \varphi \wedge \psi \Rightarrow \Gamma \vdash \varphi$ *and* $\Gamma \vdash \psi$,
(d) $\Gamma \cup \{\varphi\} \vdash \psi \Rightarrow \Gamma \vdash \varphi \to \psi$,
(e) $\Gamma \vdash \varphi, \Gamma' \vdash \varphi \to \psi \Rightarrow \Gamma \cup \Gamma' \vdash \psi$,
(f) $\Gamma \vdash \bot \Rightarrow \Gamma \vdash \varphi$,
(g) $\Gamma \cup \{\neg\varphi\} \vdash \bot \Rightarrow \Gamma \vdash \varphi$.

Proof Immediate from the definition of derivation. □

We now list some theorems. ¬ and ↔ are used as abbreviations.

Theorem 2.4.4

(1) $\vdash \varphi \to (\psi \to \varphi)$,
(2) $\vdash \varphi \to (\neg\varphi \to \psi)$,
(3) $\vdash (\varphi \to \psi) \to [(\psi \to \sigma) \to (\varphi \to \sigma)]$,
(4) $\vdash (\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi)$,
(5) $\vdash \neg\neg\varphi \leftrightarrow \varphi$,
(6) $\vdash [\varphi \to (\psi \to \sigma)] \leftrightarrow [\varphi \wedge \psi \to \sigma]$,
(7) $\vdash \bot \leftrightarrow (\varphi \wedge \neg\varphi)$.

Proof

1. $$\dfrac{\dfrac{[\varphi]^1}{\psi \to \varphi} \to I}{\varphi \to (\psi \to \varphi)} \to I_1$$

2. $$\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^2 \quad [\neg\varphi]^1}{\bot} \to E}{\psi} \bot}{\neg\varphi \to \psi} \to I_1}{\varphi \to (\neg\varphi \to \psi)} \to I_2$$

3. $$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^1 \quad [\varphi \to \psi]^3}{\psi} \to E \qquad [\psi \to \sigma]^2}{\sigma} \to E}{\varphi \to \sigma} \to I_1}{(\psi \to \sigma) \to (\varphi \to \sigma)} \to I_2}{(\varphi \to \psi) \to ((\psi \to \sigma) \to (\varphi \to \sigma))} \to I_3$$

4. For one direction, substitute $\bot$ for σ in 3, then $\vdash (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$. Conversely:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\neg\psi]^1 \quad [\neg\psi \to \neg\varphi]^3}{\neg\varphi} \to E \qquad [\varphi]^2}{\bot} \to E}{\psi}\,\mathrm{RAA}_1}{\varphi \to \psi} \to I_2}{(\neg\psi \to \neg\varphi) \to (\varphi \to \psi)} \to I_3$$

So now we have

$$\dfrac{\begin{array}{c}\mathcal{D}\\ (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)\end{array} \qquad \begin{array}{c}\mathcal{D}'\\ (\neg\psi \to \neg\varphi) \to (\varphi \to \psi)\end{array}}{(\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi)} \wedge I$$

5. We already proved $\varphi \to \neg\neg\varphi$ as an example. Conversely:

$$\dfrac{\dfrac{\dfrac{[\neg\varphi]^1 \quad [\neg\neg\varphi]^2}{\bot} \to E}{\varphi}\,\mathrm{RAA}_1}{\neg\neg\varphi \to \varphi} \to I_2$$

The result now follows. Numbers 6 and 7 are left to the reader. □

The system outlined in this section is called the "calculus of natural deduction" for a good reason: its manner of making inferences corresponds to the reasoning we intuitively use. The rules present means to take formulas apart, or to put them together. A derivation then consists of a skillful manipulation of the rules, the use of which is usually suggested by the form of the formula we want to prove.

We will discuss one example in order to illustrate the general strategy of building derivations. Let us consider the converse of our previous example **III**.

To prove $(\varphi \wedge \psi \to \sigma) \to [\varphi \to (\psi \to \sigma)]$ there is just one initial step: assume $\varphi \wedge \psi \to \sigma$ and try to derive $\varphi \to (\psi \to \sigma)$. Now we can either look at the assumption or at the desired result. Let us consider the latter one first: to show $\varphi \to (\psi \to \sigma)$, we should assume φ and derive $\psi \to \sigma$, but for the latter we should assume ψ and derive σ.

So, altogether we may assume $\varphi \wedge \psi \to \sigma$ and φ and ψ. Now the procedure suggests itself: derive $\varphi \wedge \psi$ from φ and ψ, and σ from $\varphi \wedge \psi$ and $\varphi \wedge \psi \to \sigma$.

Put together, we get the following derivation:

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^2 \quad [\psi]^1}{\varphi \wedge \psi}\wedge I \qquad [\varphi \wedge \psi \to \sigma]^3}{\sigma}\to E}{\psi \to \sigma}\to I_1}{\varphi \to (\psi \to \sigma)}\to I_2}{(\varphi \wedge \psi \to \sigma) \to (\varphi \to (\psi \to \sigma))}\to I_3$$

Had we considered $\varphi \wedge \psi \to \sigma$ first, then the only way to proceed would be to add $\varphi \wedge \psi$ and apply $\to E$. Now $\varphi \wedge \psi$ either remains an assumption, or it is obtained from something else. It immediately occurs to the reader to derive $\varphi \wedge \psi$ from φ and ψ. But now he will build up the derivation we obtained above.

Simple as this example seems, there are complications. In particular the rule of RAA is not nearly as natural as the other ones. Its use must be learned by practice; also a sense for the distinction between *constructive* and *non-constructive* will be helpful when trying to decide on when to use it.

Finally, we recall that $\top$ is an abbreviation for $\neg\bot$ (i.e. $\bot \to \bot$).

Exercises

1. Show that the following propositions are derivable:

   (a) $\varphi \to \varphi$, (d) $(\varphi \to \psi) \leftrightarrow \neg(\varphi \wedge \neg\psi)$,
   (b) $\bot \to \varphi$, (e) $(\varphi \wedge \psi) \leftrightarrow \neg(\varphi \to \neg\psi)$,
   (c) $\neg(\varphi \wedge \neg\varphi)$, (f) $\varphi \to (\psi \to (\varphi \wedge \psi))$.

2. Do the same for

   (a) $(\varphi \to \neg\varphi) \to \neg\varphi$,
   (b) $[\varphi \to (\psi \to \sigma] \leftrightarrow [\psi \to (\varphi \to \sigma)]$,
   (c) $(\varphi \to \psi) \wedge (\varphi \to \neg\psi) \to \neg\varphi$,
   (d) $(\varphi \to \psi) \to [(\varphi \to (\psi \to \sigma)) \to (\varphi \to \sigma)]$.

3. Show

   (a) $\varphi \vdash \neg(\neg\varphi \wedge \psi)$, (d) $\vdash \varphi \Rightarrow \vdash \psi \to \varphi$,
   (b) $\neg(\varphi \wedge \neg\psi), \varphi \vdash \psi$, (e) $\neg\varphi \vdash \varphi \to \psi$.
   (c) $\neg\varphi \vdash (\varphi \to \psi) \leftrightarrow \neg\varphi$,

4. Show

$$\vdash [(\varphi \to \psi) \to (\varphi \to \sigma)] \to [(\varphi \to (\psi \to \sigma))],$$
$$\vdash ((\varphi \to \psi) \to \varphi) \to \varphi.$$

5. Show

$$\Gamma \vdash \varphi \Rightarrow \Gamma \cup \Delta \vdash \varphi,$$
$$\Gamma \vdash \varphi;\ \ \Delta, \varphi \vdash \psi \Rightarrow \Gamma \cup \Delta \vdash \psi.$$

6. Give a recursive definition of the function Hyp which assigns to each derivation $\mathcal{D}$ its set of hypotheses $Hyp(\mathcal{D})$ (this is a bit stricter than the notion in Definition 2.4.2, since it is the smallest set of hypotheses, i.e. hypotheses without "garbage").
7. Analogous to the substitution operator for propositions we define a substitution operator for derivations. $\mathcal{D}[\varphi/p]$ is obtained by replacing each occurrence of p in each proposition in $\mathcal{D}$ by φ. Give a recursive definition of $\mathcal{D}[\varphi/p]$. Show that $\mathcal{D}[\varphi/p]$ is a derivation if $\mathcal{D}$ is one, and that $\Gamma \vdash \sigma \Rightarrow \Gamma[\varphi/p] \vdash \sigma[\varphi/p]$. Remark: for several purposes finer notions of substitution are required, but this one will do for us.
8. (*Substitution Theorem*) $\vdash (\varphi_1 \leftrightarrow \varphi_2) \to (\psi[\varphi_1/p] \leftrightarrow \psi[\varphi_2/p])$.
   Hint: use induction on ψ; the theorem will also follow from the Substitution Theorem for $\models$, once we have established the Completeness Theorem.
9. The *size*, $s(\mathcal{D})$, of a derivation is the number of proposition occurrences in $\mathcal{D}$. Give an inductive definition of $s(\mathcal{D})$. Show that one can prove properties of derivations by *induction on size*.
10. Give an inductive definition of the relation $\vdash$ (use the list of Lemma 2.4.3), and show that this relation coincides with the derived relation of Definition 2.4.2. Conclude that each Γ with $\Gamma \vdash \varphi$ contains a finite Δ, such that also $\Delta \vdash \varphi$.
11. Show

    (a) $\vdash \top$,
    (b) $\vdash \varphi \Leftrightarrow \vdash \varphi \leftrightarrow \top$,
    (c) $\vdash \neg\varphi \Leftrightarrow \vdash \varphi \leftrightarrow \bot$.

2.5 Completeness

In the present section we will show that "truth" and "derivability" coincide; to be precise: the relations "$\models$" and "$\vdash$" coincide. The easy part of the claim is: "derivability" implies "truth"; for derivability is established by the existence of a derivation. The latter motion is inductively defined, so we can prove the implication by induction on the derivation.

Lemma 2.5.1 (Soundness) $\Gamma \vdash \varphi \Rightarrow \Gamma \models \varphi$.

Proof Since, by Definition 2.4.2, $\Gamma \vdash \varphi$ iff there is a derivation $\mathcal{D}$ with all its hypotheses in Γ, it suffices to show: for each derivation $\mathcal{D}$ with conclusion φ and hypotheses in Γ we have $\Gamma \models \varphi$. We now use induction on $\mathcal{D}$.

(*basis*) If $\mathcal{D}$ has one element, then evidently $\varphi \in \Gamma$. The reader easily sees that $\Gamma \models \varphi$.

($\wedge$ I) Induction hypothesis: $\begin{array}{c}\mathcal{D}\\ \varphi\end{array}$ and $\begin{array}{c}\mathcal{D}'\\ \varphi'\end{array}$ are derivations and for each Γ, Γ' containing the hypotheses of $\mathcal{D}$, $\mathcal{D}'$, $\Gamma \models \varphi$, $\Gamma' \models \varphi'$.

Now let Γ'' contain the hypotheses of $\dfrac{\begin{array}{cc}\mathcal{D} & \mathcal{D}'\\ \varphi & \varphi'\end{array}}{\varphi \wedge \varphi'}$.

Choosing Γ and Γ' to be precisely the set of hypotheses of $\mathcal{D}$, $\mathcal{D}'$, we see that $\Gamma'' \supseteq \Gamma \cup \Gamma'$.

So $\Gamma'' \models \varphi$ and $\Gamma'' \models \varphi'$. Let $[\![\psi]\!]_v = 1$ for all $\psi \in \Gamma''$, then $[\![\varphi]\!]_v = [\![\varphi']\!]_v = 1$, hence $[\![\varphi \wedge \varphi']\!]_v = 1$. This shows $\Gamma'' \models \varphi \wedge \varphi'$.

($\wedge$ E) Induction hypothesis: for any Γ containing the hypotheses of $\begin{matrix}\mathcal{D}\\ \varphi\wedge\psi\end{matrix}$ we have $\Gamma \models \varphi \wedge \psi$. Consider a Γ containing all hypotheses of $\dfrac{\begin{matrix}\mathcal{D}\\ \varphi\wedge\psi\end{matrix}}{\varphi}$ and $\dfrac{\begin{matrix}\mathcal{D}\\ \varphi\wedge\psi\end{matrix}}{\psi}$. It is left to the reader to show $\Gamma \models \varphi$ and $\Gamma \models \psi$.

($\rightarrow I$) Induction hypothesis: for any Γ containing all hypotheses of $\begin{matrix}\varphi\\ \mathcal{D}\\ \psi\end{matrix}$, $\Gamma \models \psi$.

Let Γ' contain all hypotheses of $\dfrac{\begin{matrix}[\varphi]\\ \mathcal{D}\\ \psi\end{matrix}}{\varphi\rightarrow\psi}$. Now $\Gamma' \cup \{\varphi\}$ contains all hypotheses of $\begin{matrix}\varphi\\ \mathcal{D}\\ \psi\end{matrix}$, so if $[\![\varphi]\!] = 1$ and $[\![\chi]\!] = 1$ for all χ in Γ', then $[\![\psi]\!] = 1$. Therefore the truth table of $\rightarrow$ tells us that $[\![\varphi \rightarrow \psi]\!] = 1$ if all propositions in Γ' have value 1. Hence $\Gamma' \models \varphi \rightarrow \psi$.

($\rightarrow$ E) An exercise for the reader.

($\bot$) Induction hypothesis: for each Γ containing all hypotheses of $\begin{matrix}\mathcal{D}\\ \bot\end{matrix}$, $\Gamma \models \bot$. Since $[\![\bot]\!] = 0$ for all valuations, there is no valuation such that $[\![\psi]\!] = 1$ for all $\psi \in \Gamma$. Let Γ' contain all hypotheses of $\dfrac{\begin{matrix}\mathcal{D}\\ \bot\end{matrix}}{\varphi}$ and suppose that $\Gamma' \not\models \varphi$, then $[\![\psi]\!] = 1$ for all $\psi \in \Gamma'$ and $[\![\varphi]\!] = 0$ for some valuation. Since Γ' contains all hypotheses of the first derivation we have a contradiction.

(RAA) Induction hypothesis: for each Γ containing all hypotheses of $\begin{matrix}\neg\varphi\\ \mathcal{D}\\ \bot\end{matrix}$, we have $\Gamma \models \bot$. Let Γ' contain all hypotheses of $\dfrac{\begin{matrix}[\neg\varphi]\\ \mathcal{D}\\ \bot\end{matrix}}{\varphi}$ and suppose $\Gamma' \not\models \varphi$, then there exists a valuation such that $[\![\psi]\!] = 1$ for all $\psi \in \Gamma'$ and $[\![\varphi]\!] = 0$, i.e. $[\![\neg\varphi]\!] = 1$. But $\Gamma'' = \Gamma' \cup \{\neg\varphi\}$ contains all hypotheses of the first derivation and $[\![\psi]\!] = 1$ for all $\psi \in \Gamma''$. This is impossible since $\Gamma'' \models \bot$. Hence $\Gamma' \models \varphi$. □

This lemma may not seem very impressive, but it enables us to show that some propositions are not theorems, simply by showing that they are not tautologies. Without this lemma that would have been a very awkward task. We would have to show that there is no derivation (without hypotheses) of the given proposition. In general this requires insight in the nature of derivations, something which is beyond us at the moment.

Examples $\nvdash p_0$, $\nvdash (\varphi \rightarrow \psi) \rightarrow \varphi \wedge \psi$.

In the first example take the constant 0 valuation. $[\![p_0]\!] = 0$, so $\not\models p_0$ and hence $\nvdash p_0$. In the second example we are faced with a meta-proposition (a *schema*); strictly speaking it cannot be derivable (only *real* propositions can be). By $\vdash (\varphi \to \psi) \to \varphi \wedge \psi$ we mean that all propositions of that form (obtained by substituting real propositions for φ and ψ, if you like) are derivable. To refute it we need only one instance which is not derivable. Take $\varphi = \psi = p_0$.

In order to prove the converse of Lemma 2.5.1 we need a few new notions. The first one has an impressive history; it is the notion of *freedom from contradiction* or *consistency*. It was made the cornerstone of the foundations of mathematics by Hilbert.

Definition 2.5.2 A set Γ of propositions is *consistent* if $\Gamma \nvdash \perp$.

In words: one cannot derive a contradiction from Γ. The consistency of Γ can be expressed in various other forms.

Lemma 2.5.3 *The following three conditions are equivalent*:

(i) *Γ is consistent,*
(ii) *For no φ, $\Gamma \vdash \varphi$ and $\Gamma \vdash \neg\varphi$,*
(iii) *There is at least one φ such that $\Gamma \nvdash \varphi$.*

Proof Let us call Γ *inconsistent* if $\Gamma \vdash \perp$; then we can just as well prove the equivalence of

(iv) Γ is inconsistent,
(v) There is a φ such that $\Gamma \vdash \varphi$ and $\Gamma \vdash \neg\varphi$,
(vi) $\Gamma \vdash \varphi$ for all φ.
(iv) $\Rightarrow$ (vi) Let $\Gamma \vdash \perp$, i.e. there is a derivation $\mathcal{D}$ with conclusion $\perp$ and hypotheses in Γ. By ($\perp$) we can add one inference, $\perp \vdash \varphi$, to $\mathcal{D}$, so that $\Gamma \vdash \varphi$. This holds for all φ.
(vi) $\Rightarrow$ (v) Trivial.
(v) $\Rightarrow$ (iv) Let $\Gamma \vdash \varphi$ and $\Gamma \vdash \neg\varphi$. From the two associated derivations one obtains a derivation for $\Gamma \vdash \perp$ by ($\to$ E). □

Clause (vi) tells us why inconsistent sets (theories) are devoid of mathematical interest. For, if everything is derivable, we cannot distinguish between "good" and "bad" propositions. Mathematics tries to find distinctions, not to blur them.

In mathematical practice one tries to establish consistency by exhibiting a model (think of the consistency of the negation of Euclid's fifth postulate and the non-euclidean geometries). In the context of propositional logic this means looking for a suitable valuation.

Lemma 2.5.4 *If there is a valuation such that $[\![\psi]\!]_v = 1$ for all $\psi \in \Gamma$, then Γ is consistent.*

Proof Suppose $\Gamma \vdash \bot$, then by Lemma 2.5.1 $\Gamma \models \bot$, so for any valuation v $[\![(\psi)]\!]_v = 1$ for all $\psi \in \Gamma \Rightarrow [\![\bot]\!]_v = 1$. Since $[\![\bot]\!]_v = 0$ for all valuations, there is no valuation with $[\![\psi]\!]_v = 1$ for all $\psi \in \Gamma$. Contradiction. Hence Γ is consistent. □

Examples

1. $\{p_0, \neg p_1, p_1 \to p_2\}$ is consistent. A suitable valuation is one satisfying $[\![p_0]\!] = 1$, $[\![p_1]\!] = 0$.
2. $\{p_0, p_1, \ldots\}$ is consistent. Choose the constant 1 valuation.

Clause (v) of Lemma 2.5.3 tells us that $\Gamma \cup \{\varphi, \neg\varphi\}$ is inconsistent. Now, how could $\Gamma \cup \{\neg\varphi\}$ be inconsistent? It seems plausible to blame this on the derivability of φ. The following confirms this.

Lemma 2.5.5

(a) $\Gamma \cup \{\neg\varphi\}$ *is inconsistent* $\Rightarrow \Gamma \vdash \varphi$,
(b) $\Gamma \cup \{\varphi\}$ *is inconsistent* $\Rightarrow \Gamma \vdash \neg\varphi$.

Proof The assumptions of (a) and (b) yield the two derivations below: with conclusion $\bot$. By applying (RAA), and $(\to I)$, we obtain derivations with hypotheses in Γ, of φ, resp. $\neg\varphi$.

$$\begin{array}{cc} [\neg\varphi] & [\varphi] \\ \mathcal{D} & \mathcal{D}' \\ \dfrac{\bot}{\varphi}\,\text{RAA} & \dfrac{\bot}{\neg\varphi} \to I \end{array}$$

□

Definition 2.5.6 A set Γ is *maximally consistent* iff

(a) Γ is consistent,
(b) $\Gamma \subseteq \Gamma'$ and Γ' consistent $\Rightarrow \Gamma = \Gamma'$.

Remark One could replace (b) by (b′): if Γ is a proper subset of Γ', then Γ' is inconsistent. That is, by just throwing in one extra proposition, the set becomes inconsistent.

Maximally consistent sets play an important role in logic. We will show that there are lots of them.

Here is one example: $\Gamma = \{\varphi | [\![\varphi]\!] = 1\}$ for a fixed valuation. By Lemma 2.5.4 Γ is consistent. Consider a consistent set Γ' such that $\Gamma \subseteq \Gamma'$. Now let $\psi \in \Gamma'$ and suppose $[\![\psi]\!] = 0$, then $[\![\neg\psi]\!] = 1$, and so $\neg\psi \in \Gamma$.

But since $\Gamma \subseteq \Gamma'$ this implies that Γ' is inconsistent. Contradiction. Therefore $[\![\psi]\!] = 1$ for all $\psi \in \Gamma'$, so by definition $\Gamma = \Gamma'$. Moreover, from the proof of Lemma 2.5.11 it follows that this basically is the only kind of maximally consistent set we may expect.

The following fundamental lemma is proved directly. The reader may recognize in it an analogue of the maximal ideal existence lemma from ring theory (or the Boolean prime ideal theorem), which is usually proved by an application of Zorn's lemma.

Lemma 2.5.7 *Each consistent set Γ is contained in a maximally consistent set Γ^*.*

Proof There are countably many propositions, so suppose we have a list $\varphi_0, \varphi_1, \varphi_2, \dots$ of all propositions (cf. Exercise 5). We define a non-decreasing sequence of sets Γ_i such that the union is maximally consistent.

$$\Gamma_0 = \Gamma,$$
$$\Gamma_{n+1} = \begin{cases} \Gamma_n \cup \{\varphi_n\} & \text{if } \Gamma_n \cup \{\varphi_n\} \text{ is consistent,} \\ \Gamma_n & \text{else,} \end{cases}$$
$$\Gamma^* = \bigcup \{\Gamma_n \mid n \geq 0\}.$$

(a) Γ_n is consistent for all n.
Immediate, by induction on n.

(b) Γ^* is consistent.
Suppose $\Gamma^* \vdash \bot$ then, by the definition of $\bot$ there is derivation $\mathcal{D}$ of $\bot$ with hypotheses in Γ^*; $\mathcal{D}$ has finitely many hypotheses $\psi_0, \dots, \psi_k$. Since $\Gamma^* = \bigcup\{\Gamma_n | n \geq 0\}$, we have for each $i \leq k$ $\psi_i \in \Gamma_{n_i}$ for some n_i. Let n be $\max\{n_i | i \leq k\}$, then $\psi_0, \dots, \psi_k \in \Gamma_n$ and hence $\Gamma_n \vdash \bot$. But Γ_n is consistent. Contradiction.

(c) Γ^* is maximally consistent. Let $\Gamma^* \subseteq \Delta$ and Δ consistent. If $\psi \in \Delta$, then $\psi = \varphi_m$ for some m. Since $\Gamma_m \subseteq \Gamma^* \subseteq \Delta$ and Δ is consistent, $\Gamma_m \cup \{\varphi_m\}$ is consistent. Therefore $\Gamma_{m+1} = \Gamma_m \cup \{\varphi_m\}$, i.e. $\varphi_m \in \Gamma_{m+1} \subseteq \Gamma^*$. This shows $\Gamma^* = \Delta$.

Lemma 2.5.8 *If Γ is maximally consistent, then Γ is closed under derivability* (*i.e.* $\Gamma \vdash \varphi \Rightarrow \varphi \in \Gamma$).

Proof Let $\Gamma \vdash \varphi$ and suppose $\varphi \notin \Gamma$. Then $\Gamma \cup \{\varphi\}$ must be inconsistent. Hence $\Gamma \vdash \neg\varphi$, so Γ is inconsistent. Contradiction. □

Lemma 2.5.9 *Let Γ be maximally consistent*; *then*

for all φ *either* $\varphi \in \Gamma$, *or* $\neg\varphi \in \Gamma$,

for all φ, ψ $\varphi \to \psi \in \Gamma \Leftrightarrow (\varphi \in \Gamma \Rightarrow \psi \in \Gamma)$.

Proof (a) We know that not both φ and $\neg\varphi$ can belong to Γ. Consider $\Gamma' = \Gamma \cup \{\varphi\}$. If Γ' is inconsistent, then, by Lemmas 2.5.5, 2.5.8, $\neg\varphi \in \Gamma$. If Γ' is consistent, then $\varphi \in \Gamma$ by the maximality of Γ.

(b) Let $\varphi \to \psi \in \Gamma$ and $\varphi \in \Gamma$. To show: $\psi \in \Gamma$. Since $\varphi, \varphi \to \psi \in \Gamma$ and since Γ is closed under derivability (Lemma 2.5.8), we get $\psi \in \Gamma$ by $\to E$.

Conversely: let $\varphi \in \Gamma \Rightarrow \psi \in \Gamma$. If $\varphi \in \Gamma$ then obviously $\Gamma \vdash \psi$, so $\Gamma \vdash \varphi \to \psi$. If $\varphi \notin \Gamma$, then $\neg\varphi \in \Gamma$, and hence $\Gamma \vdash \neg\varphi$. Therefore $\Gamma \vdash \varphi \to \psi$. □

Note that we automatically get the following.

Corollary 2.5.10 *If Γ is maximally consistent, then $\varphi \in \Gamma \Leftrightarrow \neg\varphi \notin \Gamma$, and $\neg\varphi \in \Gamma \Leftrightarrow \varphi \notin \Gamma$.*

Lemma 2.5.11 *If Γ is consistent, then there exists a valuation such that $[\![\psi]\!] = 1$ for all $\psi \in \Gamma$.*

Proof (a) By Lemma 2.5.7 Γ is contained in a maximally consistent Γ^*.

(b) Define $v(p_i) = \begin{cases} 1 \text{ if } p_i \in \Gamma^* \\ 0 \text{ else} \end{cases}$ and extend v to the valuation $[\![\ ]\!]_v$.

Claim: $[\![\varphi]\!] = 1 \Leftrightarrow \varphi \in \Gamma^*$. Use induction on φ.

1. For atomic φ the claim holds by definition.
2. $\varphi = \psi \wedge \sigma$. $[\![\varphi]\!]_v = 1 \Leftrightarrow [\![\psi]\!]_v = [\![\sigma]\!]_v = 1 \Leftrightarrow$ (induction hypothesis) $\psi, \sigma \in \Gamma^*$ and so $\varphi \in \Gamma^*$. Conversely $\psi \wedge \sigma \in \Gamma^* \Leftrightarrow \psi, \sigma \in \Gamma^*$ (Lemma 2.5.8). The rest follows from the induction hypothesis.
3. $\varphi = \psi \to \sigma$. $[\![\psi \to \sigma]\!]_v = 0 \Leftrightarrow [\![\psi]\!]_v = 1$ and $[\![\sigma]\!]_v = 0 \Leftrightarrow$ (induction hypothesis) $\psi \in \Gamma^*$ and $\sigma \notin \Gamma^* \Leftrightarrow \psi \to \sigma \notin \Gamma^*$ (by Lemma 2.5.9).

(c) Since $\Gamma \subseteq \Gamma^*$ we have $[\![\psi]\!]_v = 1$ for all $\psi \in \Gamma$. □

Corollary 2.5.12 $\Gamma \nvdash \varphi \Leftrightarrow$ *there is a valuation such that $[\![\psi]\!] = 1$ for all $\psi \in \Gamma$ and $[\![\varphi]\!] = 0$.*

Proof $\Gamma \nvdash \varphi \Leftrightarrow \Gamma \cup \{\neg\varphi\}$ consistent $\Leftrightarrow$ there is a valuation such that $[\![\psi]\!] = 1$ for all $\psi \in \Gamma \cup \{\neg\varphi\}$, or $[\![\psi]\!] = 1$ for all $\psi \in \Gamma$ and $[\![\varphi]\!] = 0$. □

Theorem 2.5.13 (Completeness Theorem) $\Gamma \vdash \varphi \Leftrightarrow \Gamma \models \varphi$.

Proof $\Gamma \nvdash \varphi \Rightarrow \Gamma \not\models \varphi$ by Corollary 2.5.12. The converse holds by Lemma 2.5.1. □

In particular we have $\vdash \varphi \Leftrightarrow\ \models \varphi$, so the set of theorems is exactly the set to tautologies.

The Completeness Theorem tells us that the tedious task of making derivations can be replaced by the (equally tedious, but automatic) task of checking tautologies. This simplifies, at least in theory, the search for theorems considerably; for derivations one has to be (moderately) clever, for truth tables one has to possess perseverance.

For logical theories one sometimes considers another notion of completeness: a set Γ is called *complete* if for each φ, either $\Gamma \vdash \varphi$, or $\Gamma \vdash \neg\varphi$. This notion is closely related to "maximally consistent". From Exercise 6 it follows that

$\mathrm{Cons}(\Gamma) = \{\sigma \mid \Gamma \vdash \sigma\}$ (*the set of consequences of* Γ) is maximally consistent if Γ is a complete set. The converse also holds (cf. Exercise 10). Propositional logic itself (i.e. the case $\Gamma = \emptyset$) is not complete in this sense, e.g. $\nvdash p_0$ and $\nvdash \neg p_0$.

There is another important notion which is traditionally considered in logic: that of *decidability*. Propositional logic is decidable in the following sense: there is an effective procedure to check the derivability of propositions φ. Stated otherwise: there is an algorithm that for each φ tests if $\vdash \varphi$.

The algorithm is simple: write down the complete truth table for φ and check if the last column contains only 1's. If so, then $\models \varphi$ and, by the Completeness Theorem, $\vdash \varphi$. If not, then $\not\models \varphi$ and hence $\nvdash \varphi$. This is certainly not the best possible algorithm, one can find more economical ones. There are also algorithms that give more information, e.g. they not only test $\vdash \varphi$, but also yield a derivation, if one exists. Such algorithms require, however, a deeper analysis of derivations, which falls outside the scope of this book.

There is one aspect of the Completeness Theorem that we want to discuss now. It does not come as a surprise that truth follows from derivability. After all we start with a combinatorial notion, defined inductively, and we end up with "being true for all valuations". A simple inductive proof does the trick.

For the converse the situation is totally different. By definition $\Gamma \models \varphi$ means that $[\![\varphi]\!]_v = 1$ for all valuations v that make all propositions of Γ true. So we know something about the behavior of *all* valuations with respect to Γ and φ. Can we hope to extract from such infinitely many set theoretical facts the finite, concrete information needed to build a derivation for $\Gamma \vdash \varphi$? Evidently the available facts do not give us much to go on. Let us therefore simplify matters a bit by cutting down the Γ; after all we use only finitely many formulas of Γ in a derivation, so let us suppose that those formulas $\psi_1, \dots, \psi_n$ are given. Now we can hope for more success, since only finitely many atoms are involved, and hence we can consider a finite "part" of the infinitely many valuations that play a role. That is, only the restrictions of the valuations to the set of atoms occurring in $\psi_1, \dots, \psi_n, \varphi$ are relevant. Let us simplify the problem one more step. We know that $\psi_1, \dots, \psi_n \vdash \varphi$ ($\psi_1, \dots, \psi_n \models \varphi$) can be replaced by $\vdash \psi_1 \wedge \dots \wedge \psi_n \to \varphi (\models \psi_1 \wedge \dots \wedge \psi_n \to \varphi)$, on the ground of the derivation rules (the definition of valuation). So we ask ourselves: given the truth table for a tautology σ, can we effectively find a derivation for σ? This question is not answered by the Completeness Theorem, since our proof of it is not effective (at least not prima facie so). It has been answered positively, e.g. by Post, Bernays and Kalmar (cf. Kleene 1952, IV, §29) and it is easily treated by means of Gentzen techniques, or semantic tableaux. We will just sketch a method of proof: we can effectively find a conjunctive normal form σ^* for σ such that $\vdash \sigma \leftrightarrow \sigma^*$. It is easily shown that σ^* is a tautology iff each conjunct contains an atom and its negation, or $\neg \perp$, and glue it all together to obtain a derivation of σ^*, which immediately yields a derivation of σ.

Exercises

1. Check which of the following sets are consistent:
   (a) $\{\neg p_1 \wedge p_2 \to p_0, p_1 \to (\neg p_1 \to p_2), p_0 \leftrightarrow \neg p_2\}$,

(b) $\{p_0 \to p_1, p_1 \to p_2, p_2 \to p_3, p_3 \to \neg p_0\}$,
(c) $\{p_0 \to p_1, p_0 \wedge p_2 \to p_1 \wedge p_3, p_0 \wedge p_2 \wedge p_4 \to p_1 \wedge p_3 \wedge p_5, \ldots\}$.

2. Show that the following are equivalent:
   (a) $\{\varphi_1, \ldots, \varphi_n\}$ is consistent.
   (b) $\nvdash \neg(\varphi_1 \wedge \varphi_2 \wedge \cdots \wedge \varphi_n)$.
   (c) $\nvdash \varphi_1 \wedge \varphi_2 \wedge \cdots \wedge \varphi_{n-1} \to \neg\varphi_n$.
3. φ is *independent* from Γ if $\Gamma \nvdash \varphi$ and $\Gamma \nvdash \neg\varphi$. Show that: $p_1 \to p_2$ is independent from $\{p_1 \leftrightarrow p_0 \wedge \neg p_2, p_2 \to p_0\}$.
4. A set Γ is *independent* if for each $\varphi \in \Gamma$ $\Gamma - \{\varphi\} \nvdash \varphi$.
   (a) Show that each finite set Γ has an independent subset Δ such that $\Delta \vdash \varphi$ for all $\varphi \in \Gamma$.
   (b) Let $\Gamma = \{\varphi_0, \varphi_1, \varphi_2, \ldots\}$. Find an equivalent set $\Gamma' = \{\psi_0, \psi_1, \ldots\}$ (i.e. $\Gamma \vdash \psi_i$ and $\Gamma' \vdash \varphi_i$ for all i) such that $\vdash \psi_{n+1} \to \psi_n$, but $\nvdash \psi_n \to \psi_{n+1}$. Note that Γ' may be finite.
   (c) Consider an infinite Γ' as in (b). Define $\sigma_0 = \psi_0, \sigma_{n+1} = \psi_n \to \psi_{n+1}$. Show that $\Delta = \{\sigma_0, \sigma_1, \sigma_2, \ldots\}$ is independent and equivalent to Γ'.
   (d) Show that each set Γ is equivalent to an independent set Δ.
   (e) Show that Δ need not be a subset of Γ (consider $\{p_0, p_0 \wedge p_1, p_0 \wedge p_1 \wedge p_2, \ldots\}$).
5. Find an effective way of enumerating all propositions (hint: consider sets Γ_n of all propositions of rank $\le n$ with atoms from $p_0, \ldots, p_n$).
6. Show that a consistent set Γ is maximally consistent if either $\varphi \in \Gamma$ or $\neg\varphi \in \Gamma$ for all φ.
7. Show that $\{p_0, p_1, p_2, \ldots, p_n, \ldots\}$ is complete.
8. (*Compactness Theorem*). Show: there is a v such that $[\![\psi]\!]_v = 1$ for all $\psi \in \Gamma \Leftrightarrow$ for each finite subset $\Delta \subseteq \Gamma$ there is a v such that $[\![\sigma]\!]_v = 1$ for all $\sigma \in \Delta$.
   Formulated in terms of Exercise 13 of 2.3: $[\![\Gamma]\!] \neq \emptyset$ if $[\![\Delta]\!] \neq \emptyset$ for all finite $\Delta \subseteq \Gamma$.
9. Consider an infinite set $\{\varphi_1, \varphi_2, \varphi_3, \ldots\}$. If for each valuation there is an n such that $[\![\varphi_n]\!] = 1$, then there is an m such that $\vdash \varphi_1 \vee \cdots \vee \varphi_m$. (Hint: consider the negations $\neg\varphi_1, \neg\varphi_2 \ldots$ and apply Exercise 8.)
10. Show: $\mathrm{Cons}(\Gamma) = \{\sigma \mid \Gamma \vdash \sigma\}$ is maximally consistent $\Leftrightarrow$ Γ is complete.
11. Show: Γ is maximally consistent $\Leftrightarrow$ there is a unique valuation such that $[\![\psi]\!] = 1$ for all $\psi \in \Gamma$, where Γ is a theory, i.e. Γ is closed under $\vdash$ ($\Gamma \vdash \sigma \Rightarrow \sigma \in \Gamma$).
12. Let φ be a proposition containing the atom p. For convenience we write $\varphi(\sigma)$ for $\varphi[\sigma/p]$.
    As before we abbreviate $\neg\bot$ by $\top$.
    Show:
    (i) $\varphi(\top) \vdash \varphi(\top) \leftrightarrow \top$ and $\varphi(\top) \vdash \varphi(\varphi(\top))$.
    (ii) $\neg\varphi(\top) \vdash \varphi(\top) \leftrightarrow \bot$,
    $\varphi(p), \neg\varphi(\top) \vdash p \leftrightarrow \bot$,
    $\varphi(p), \neg\varphi(\top) \vdash \varphi(\varphi(\top))$.
    (iii) $\varphi(p) \vdash \varphi(\varphi(\top))$.

13. If the atoms p and q do not occur in ψ and φ respectively, then

$$\models \varphi(p) \to \psi \Rightarrow \models \varphi(\sigma) \to \psi \text{ for all } \sigma,$$
$$\models \varphi \to \psi(q) \Rightarrow \models \varphi \to \psi(\sigma) \text{ for all } \sigma.$$

14. Let $\vdash \varphi \to \psi$. We call σ an *interpolant* if $\vdash \varphi \to \sigma$ and $\vdash \sigma \to \psi$, and moreover σ contains only atoms common to φ and ψ. Consider $\varphi(p, r), \psi(r, q)$ with all atoms displayed. Show that $\varphi(\varphi(\top, r), r)$ is an interpolant (use Exercises 12, 13).
15. Prove the general *interpolation theorem* (Craig): For any φ, ψ with $\vdash \varphi \to \psi$ there exists an interpolant (iterate the procedure of Exercise 13).

2.6 The Missing Connectives

The language of Sect. 2.4 contained only the connectives $\wedge, \to$ and $\bot$. We already know that, from the semantical point of view, this language is sufficiently rich, i.e. the missing connectives can be defined. As a matter of fact we have already used the negation as a defined notion in the preceding sections.

It is a matter of sound mathematical practice to introduce new notions if their use simplifies our labor, and if they codify informal existing practice. This, clearly, is a reason for introducing $\neg, \leftrightarrow$ and $\vee$.

Now there are two ways to proceed: one can introduce the new connectives as abbreviations (of complicated propositions), or one can enrich the language by actually adding the connectives to the alphabet, and providing rules of derivation.

The first procedure was adopted above; it is completely harmless, e.g. each time one reads $\varphi \leftrightarrow \psi$, one has to replace it by $(\varphi \to \psi) \wedge (\psi \to \varphi)$. So it is nothing but a shorthand, introduced for convenience. The second procedure is of a more theoretical nature. The language is enriched and the set of derivations is enlarged. As a consequence one has to review the theoretical results (such as the Completeness Theorem) obtained for the simpler language.

We will adopt the first procedure and also outline the second approach.

Definition 2.6.1

$$\begin{aligned} \varphi \vee \psi &:= \neg(\neg\varphi \wedge \neg\psi), \\ \neg\varphi &:= \varphi \to \bot, \\ \varphi \leftrightarrow \psi &:= (\varphi \to \psi) \wedge (\psi \to \varphi). \end{aligned}$$

N.B. This means that the above expressions are *not* part of the language, but abbreviations for certain propositions.

The properties of $\vee, \neg$ and $\leftrightarrow$ are given in the following lemma.

Lemma 2.6.2

(i) $\varphi \vdash \varphi \vee \psi$, $\psi \vdash \varphi \vee \psi$,
(ii) $\Gamma, \varphi \vdash \sigma$ *and* $\Gamma, \psi \vdash \sigma \Rightarrow \Gamma, \varphi \vee \psi \vdash \sigma$,
(iii) $\varphi, \neg\varphi \vdash \bot$,
(iv) $\Gamma, \varphi \vdash \bot \Rightarrow \Gamma \vdash \neg\varphi$,
(v) $\varphi \leftrightarrow \psi, \varphi \vdash \psi$ *and* $\varphi \leftrightarrow \psi, \psi \vdash \varphi$,
(vi) $\Gamma, \varphi \vdash \psi$ *and* $\Gamma, \psi \vdash \varphi \Rightarrow \Gamma \vdash \varphi \leftrightarrow \psi$.

Proof The only non-trivial part is (*ii*). We exhibit a derivation of σ from Γ and $\varphi \vee \psi$ (i.e. $\neg(\neg\varphi \wedge \neg\psi)$), given derivations $\mathcal{D}_1$ and $\mathcal{D}_2$ of $\Gamma, \varphi \vdash \sigma$ and $\Gamma, \psi \vdash \sigma$.

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\begin{matrix}[\varphi]^1\\ \mathcal{D}_1\\ \sigma\end{matrix} \quad [\neg\sigma]^3}{\bot}\rightarrow E}{\neg\varphi}\rightarrow I_1 \qquad \dfrac{\dfrac{\begin{matrix}[\psi]^2\\ \mathcal{D}_2\\ \sigma\end{matrix} \quad [\neg\sigma]^3}{\bot}\rightarrow E}{\neg\psi}\rightarrow I_2}{\neg\varphi \wedge \neg\psi}\wedge I \qquad \neg(\neg\varphi \wedge \neg\psi)}{\bot}\rightarrow E}{\sigma}\text{RAA}_3$$

The remaining cases are left to the reader. □

Note that (i) and (ii) read as introduction and elimination rules for $\vee$, (iii) and (iv) as ditto for $\neg$, (vi) and (v) as ditto for $\leftrightarrow$.

They legalize the following shortcuts in derivations:

$$\dfrac{\varphi}{\varphi \vee \psi}\vee I \qquad \dfrac{\psi}{\varphi \vee \psi}\vee I \qquad \dfrac{\varphi \vee \psi \quad \begin{matrix}[\varphi]\\ \vdots\\ \sigma\end{matrix} \quad \begin{matrix}[\psi]\\ \vdots\\ \sigma\end{matrix}}{\sigma}\vee E$$

$$\dfrac{\begin{matrix}[\varphi]\\ \vdots\\ \bot\end{matrix}}{\neg\varphi}\neg I \qquad\qquad \dfrac{\varphi \quad \neg\varphi}{\bot}\neg E$$

$$\dfrac{\begin{matrix}[\varphi] & [\psi] \\ \vdots & \vdots \\ \psi & \varphi\end{matrix}}{\varphi \leftrightarrow \psi} \leftrightarrow I \qquad\qquad \dfrac{\varphi \quad \varphi\leftrightarrow\psi}{\psi} \qquad \dfrac{\psi \quad \varphi\leftrightarrow\psi}{\varphi} \leftrightarrow E$$

Consider for example an application of $\vee E$

$$\dfrac{\begin{matrix} & [\varphi] & [\psi] \\ \mathcal{D}_0 & \mathcal{D}_1 & \mathcal{D}_2 \\ \varphi\vee\psi & \sigma & \sigma\end{matrix}}{\sigma}\vee E$$

This is a mere shorthand for

$$\dfrac{\dfrac{\begin{matrix}\mathcal{D}_0 \\ \neg(\neg\varphi\wedge\neg\psi)\end{matrix} \qquad \dfrac{\dfrac{\dfrac{\begin{matrix}[\varphi]^1 \\ \mathcal{D}_1 \\ \sigma\end{matrix} \quad [\neg\sigma]^3}{\bot}}{\neg\varphi}\,1 \qquad \dfrac{\dfrac{\begin{matrix}[\psi]^2 \\ \mathcal{D}_2 \\ \sigma\end{matrix} \quad [\neg\sigma]^3}{\bot}}{\neg\psi}\,2}{\neg\varphi\wedge\neg\psi}}{\bot}\,1}{\sigma}\,3$$

The reader is urged to use the above shortcuts in actual derivations, whenever convenient. As a rule, only $\vee I$ and $\vee E$ are of importance; the reader has of course recognized the rules for $\neg$ and $\leftrightarrow$ as slightly eccentric applications of familiar rules.

Examples $\vdash (\varphi\wedge\psi)\vee\sigma \leftrightarrow (\varphi\vee\sigma)\wedge(\psi\vee\sigma)$.

$$\dfrac{\dfrac{(\varphi\wedge\psi)\vee\sigma \qquad \dfrac{\dfrac{[\varphi\wedge\psi]^1}{\varphi}}{\varphi\vee\sigma} \qquad \dfrac{[\sigma]^1}{\varphi\vee\sigma}}{\varphi\vee\sigma}\,1 \qquad \dfrac{(\varphi\wedge\psi)\vee\sigma \qquad \dfrac{\dfrac{[\varphi\wedge\psi]^2}{\psi}}{\psi\vee\sigma} \qquad \dfrac{[\sigma]^2}{\psi\vee\sigma}}{\psi\vee\sigma}\,2}{(\varphi\vee\sigma)\wedge(\psi\vee\sigma)} \tag{2.2}$$

Conversely

$$\dfrac{\dfrac{(\varphi \vee \sigma) \wedge (\psi \vee \sigma)}{\varphi \vee \sigma} \qquad \dfrac{\dfrac{(\varphi \vee \sigma) \wedge (\psi \vee \sigma)}{\psi \vee \sigma} \qquad \dfrac{\dfrac{[\varphi]^2 \quad [\psi]^1}{\varphi \wedge \psi}}{(\varphi \wedge \psi) \vee \sigma} \qquad \dfrac{[\sigma]^1}{(\varphi \wedge \psi) \vee \sigma}}{(\varphi \wedge \psi) \vee \sigma}\,1 \qquad \dfrac{[\sigma]^2}{(\varphi \wedge \psi) \vee \sigma}}{(\varphi \wedge \psi) \vee \sigma}\,2 \tag{2.3}$$

Combining (2.2) and (2.3) we get one derivation:

$$\dfrac{\begin{matrix}[(\varphi \wedge \psi) \vee \sigma] \\ \mathcal{D} \\ (\varphi \vee \sigma) \wedge (\psi \vee \sigma)\end{matrix} \qquad \begin{matrix}[(\varphi \vee \sigma) \wedge (\psi \vee \sigma)] \\ \mathcal{D}' \\ (\varphi \wedge \psi) \vee \sigma\end{matrix}}{(\varphi \wedge \psi) \vee \sigma \leftrightarrow (\varphi \vee \sigma) \wedge (\psi \vee \sigma)} \leftrightarrow I$$

$\vdash \varphi \vee \neg\varphi$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^1}{\varphi \vee \neg\varphi} \vee I \qquad [\neg(\varphi \vee \neg\varphi)]^2}{\bot} \to E}{\neg\varphi} \to I_1}{\varphi \vee \neg\varphi} \vee I \qquad [\neg(\varphi \vee \neg\varphi)]^2}{\bot} \to E}{\varphi \vee \neg\varphi} \text{RAA}_2$$

$\vdash (\varphi \to \psi) \vee (\psi \to \varphi)$

$$\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{\dfrac{[\varphi]^1}{\psi \to \varphi} \to I_1}{(\varphi \to \psi) \vee (\psi \to \varphi)} \vee I \qquad [\neg((\varphi \to \psi) \vee (\psi \to \varphi))]^2}{\bot} \to E}{\psi} \bot}{\varphi \to \psi} \to I_1}{(\varphi \to \psi) \vee (\psi \to \varphi)} \vee I \qquad [\neg((\varphi \to \psi) \vee (\psi \to \varphi))]^2}{\bot} \to E}{(\varphi \to \psi) \vee (\psi \to \varphi)} \text{RAA}_2$$

$\vdash \neg(\varphi \wedge \psi) \to \neg\varphi \vee \neg\psi$

$$\dfrac{\dfrac{\dfrac{[\neg(\varphi \wedge \psi)] \qquad \dfrac{\dfrac{\dfrac{[\neg(\neg\varphi \vee \neg\psi)] \quad \dfrac{[\neg\varphi]}{\neg\varphi \vee \neg\psi}}{\bot}}{\varphi} \qquad \dfrac{\dfrac{[\neg(\neg\varphi \vee \neg\psi)] \quad \dfrac{[\neg\psi]}{\neg\varphi \vee \neg\psi}}{\bot}}{\psi}}{\varphi \wedge \psi}}{\bot}}{\neg\varphi \vee \neg\psi}}{\neg(\varphi \wedge \psi) \to \neg\varphi \vee \neg\psi}$$

We now give a sketch of the second approach. We add $\vee$, $\neg$ and $\leftrightarrow$ to the language, and extend the set of propositions correspondingly. Next we add the rules for $\vee$, $\neg$ and $\leftrightarrow$ listed above to our stock of derivation rules. To be precise we should now also introduce a new derivability sign. However, we will stick to the trusted $\vdash$ in the expectation that the reader will remember that now we are making derivations in a larger system. The following holds.

Theorem 2.6.3

$$\vdash \varphi \vee \psi \leftrightarrow \neg(\neg\varphi \wedge \neg\psi).$$
$$\vdash \neg\varphi \leftrightarrow (\varphi \to \bot).$$
$$\vdash (\varphi \leftrightarrow \psi) \leftrightarrow (\varphi \to \psi) \wedge (\psi \to \varphi).$$

Proof Observe that by Lemma 2.6.2 the defined and the primitive (real) connectives obey exactly the same derivability relations (derivation rules, if you wish). This leads immediately to the desired result. Let us give one example.

$\varphi \vdash \neg(\neg\varphi \wedge \neg\psi)$ and $\psi \vdash \neg(\neg\varphi \wedge \neg\psi)$ (2.6.2 (i)), so by $\vee E$ we get

$$\varphi \vee \psi \vdash \neg(\neg\varphi \wedge \neg\psi) \ldots \quad (1)$$

Conversely $\varphi \vdash \varphi \vee \psi$ and $\psi \vdash \varphi \vee \psi$ (by $\vee I$), hence by 2.6.2 (ii)

$$\neg(\neg\varphi \wedge \neg\psi) \vdash \varphi \vee \psi \ldots \quad (2)$$

Apply $\leftrightarrow I$, to (1) and (2), then $\vdash \varphi \vee \psi \leftrightarrow \neg(\neg\varphi \wedge \neg\psi)$. The rest is left to the reader. □

For more results the reader is directed to the exercises.

The rules for $\vee$, $\leftrightarrow$, and $\neg$ indeed capture the intuitive meaning of those connectives. Let us consider disjunction: $(\vee\mathrm{I})$. If we know φ then we certainly know $\varphi \vee \psi$ (we even know exactly which disjunct). The rule $(\vee\mathrm{E})$ captures the idea of "proof by cases": if we know $\varphi \vee \psi$ and in each of both cases we can conclude σ, then we may outright conclude σ. Disjunction intuitively calls for a decision: which of the two disjuncts is given or may be assumed? This constructive streak of $\vee$ is crudely but conveniently blotted out by the identification of $\varphi \vee \psi$ and $\neg(\neg\varphi \wedge \neg\psi)$. The latter only tells us that φ and ψ cannot both be wrong, but not which one is right. For more information on this matter of constructiveness, which plays a role in demarcating the borderline between two-valued classical logic and effective intuitionistic logic, the reader is referred to Chap. 6.

Note that with $\vee$ as a primitive connective some theorems become harder to prove. For example, $\vdash \neg(\neg\neg\varphi \wedge \neg\varphi)$ is trivial, but $\vdash \varphi \vee \neg\varphi$ is not. The following rule of thumb may be useful: going from non-effective (or no) premises to an effective conclusion calls for an application of *RAA*.

Exercises

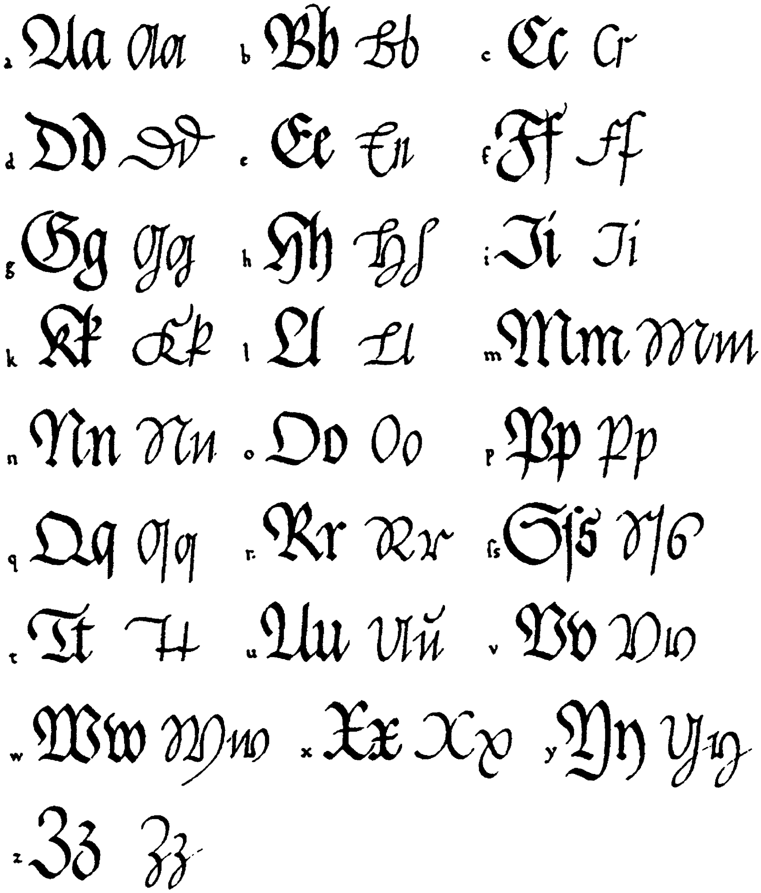
1. Show $\vdash \varphi \vee \psi \to \psi \vee \varphi, \vdash \varphi \vee \varphi \leftrightarrow \varphi$.
2. Consider the full language $\mathcal{L}$ with the connectives $\wedge, \to, \bot, \leftrightarrow \vee$ and the restricted language $\mathcal{L}'$ with connectives $\wedge, \to, \bot$. Using the appropriate derivation rules we get the derivability notions $\vdash$ and $\vdash'$. We define an obvious translation from $\mathcal{L}$ into $\mathcal{L}'$:

$$\begin{aligned}
\varphi^+ &:= \varphi \quad \text{for atomic } \varphi \\
(\varphi \Box \psi)^+ &:= \varphi^+ \Box \psi^+ \quad \text{for } \Box = \wedge, \to, \\
(\varphi \vee \psi)^+ &:= \neg(\neg\varphi^+ \wedge \neg\varphi^+), \quad \text{where } \neg \text{ is an abbreviation,} \\
(\varphi \leftrightarrow \psi)^+ &:= (\varphi^+ \to \psi^+) \wedge (\psi^+ \to \varphi^+), \\
(\neg\varphi)^+ &:= \varphi^+ \to \bot .
\end{aligned}$$

   Show
   (i) $\vdash \varphi \leftrightarrow \varphi^+$,
   (ii) $\vdash \varphi \Leftrightarrow \vdash' \varphi^+$,
   (iii) $\varphi^+ = \varphi$ for $\varphi \in \mathcal{L}'$.
   (iv) Show that the full logic is *conservative* over the restricted logic, i.e. for $\varphi \in \mathcal{L}'$ $\vdash \varphi \Leftrightarrow \vdash' \varphi$.
3. Show that the Completeness Theorem holds for the full logic. Hint: use Exercise 2.
4. Show
   (a) $\vdash \top \vee \bot$.
   (b) $\vdash (\varphi \leftrightarrow \top) \vee (\varphi \leftrightarrow \bot)$.
   (c) $\vdash \varphi \leftrightarrow (\varphi \leftrightarrow \top)$.
5. Show $\vdash (\varphi \vee \psi) \leftrightarrow ((\varphi \to \psi) \to \psi)$.
6. Show
   (a) Γ is complete $\Leftrightarrow$ ($\Gamma \vdash \varphi \vee \psi \Leftrightarrow \Gamma \vdash \varphi$ or $\Gamma \vdash \psi$, for all φ, ψ),
   (b) Γ is maximally consistent $\Leftrightarrow$ Γ is a consistent theory and for all φ, ψ ($\varphi \vee \psi \in \Gamma \Leftrightarrow \varphi \in \Gamma$ or $\psi \in \Gamma$).
7. Show in the system with $\vee$ as a primitive connective

$$\vdash (\varphi \to \psi) \leftrightarrow (\neg\varphi \vee \psi),$$
$$\vdash (\varphi \to \psi) \vee (\psi \to \varphi).$$

Gothic Alphabet



http://www.springer.com/978-1-4471-4557-8

Logic and Structure
van Dalen, D.
2013, X, 263 p. 27 illus., Softcover
ISBN: 978-1-4471-4557-8