Chapter 2
Propositional Logic

H.C.M. (Harrie) de Swart

Abstract In this chapter we analyse reasoning patterns of which the validity only depends on the meaning of the propositional connectives 'if ..., then ...', 'and', 'or' and 'not'. By giving a precise description of the meaning of these propositional connectives one is able to give a precise definition of the notion of logical or valid consequence. Two such definitions are given: a semantic one, in terms of truth values and hence in terms of the meaning of the formulas involved, and a syntactic one in terms of logical axioms and rules of which only the form is important. The semantic and the syntactic definition of logical consequence turn out be equivalent, giving us confidence that we gave a proper characterization of the intuitive notion of logical consequence. We prove or disprove all kinds of statements *about* the notion of logical or valid consequence, which is useful in order to get a good grasp of this notion. The last section treats a number of paradoxes which have been important for the progress in science and philosophy; it also contains a number of historical and philosophical remarks.

2.1 Linguistic Considerations

Logic is such a rich, broad and varied discipline that it is necessary to approach it by picking a small and manageable portion to treat first, after which the treatment can be extended to include more. In this Chapter we restrict our study of reasoning to what is called *propositional logic* or the *propositional calculus*.

A *proposition* is the meaning of a declarative sentence, like 'John is ill', 'Coby goes to school', etc., where a sentence has been obtained from letters or words from a given alphabet according to certain grammatical rules. So, a sentence is just a combination of letters or words, while the corresponding proposition is the meaning of the sentence in question. One says that a sentence *expresses* a proposition. This explains the term '*sentential calculus*' instead of 'propositional calculus'. A proposition is either true or false, although we do not have to know which of the two.

H. de Swart, *Philosophical and Mathematical Logic*, Springer Undergraduate Texts in Philosophy, https://doi.org/10.1007/978-3-030-03255-5_2

Besides declarative sentences one can distinguish interrogatory sentences which ask questions and imperative sentences which express commands. These latter sentences do not express propositions: it does not make sense to ask whether they express something true or something false. Note that different declarative sentences may express the same proposition. Thus the same proposition is expressed by 'John reads the book' and 'the book is read by John'. $3^2+4^2=5^2$ and $5^2=3^2+4^2$ also express the same proposition (which happens to be true); but $5+7=12$ expresses a different proposition (also true).

By means of connectives one may construct more complex propositions from more elementary ones. For instance, 'John is ill and Coby goes to school' has been composed from the two more elementary propositions by means of the connective 'and'. The most important *connectives* or *propositional operations* are: 'if and only if (iff)', 'if ..., then ...', 'and', 'or' and 'not'. In propositional logic one uses the symbols $\rightleftarrows$, $\rightarrow$, $\wedge$, $\vee$ and $\neg$ for these connectives, respectively.

We distinguish *atomic propositions*, like 'John is ill' and 'Coby goes to school' on the one hand and *composite propositions* on the other hand. *Atomic propositions* are those propositions which cannot be composed of yet more simple propositions by means of propositional operations. If a proposition has been composed from more elementary propositions by means of one or more propositional operations we call it a composite proposition. Thus, 'John is ill and Coby goes to school' is a composite proposition.

In propositional logic one uses letters $P_1, P_2, P_3, \ldots$ to denote atomic propositions. For instance, 'John is ill' may be translated by P_1, while 'Coby goes to school' may be translated by P_2. The composite proposition 'John is ill and Coby goes to school' is then translated by $P_1 \wedge P_2$.

In *propositional logic* one studies the (in)validity of reasoning patterns of which the (in)validity is completely determined by the meaning of the connectives 'if and only if' ($\rightleftarrows$), 'if ..., then ...' ($\rightarrow$), 'and' ($\wedge$), 'or' ($\vee$) and 'not' ($\neg$) between the propositions in question. A simple example is the following reasoning pattern, called *Modus Ponens* (MP):

It snows (1)	P_1
If it snows, then it is cold. (2)	$P_1 \rightarrow P_2$
Therefore: it is cold. (3)	P_2

The pattern to the above right is *valid*, i.e., no matter what propositions the formulas P_1, P_2 stand for, if the resulting two premisses are both true, then also the conclusion must be true; in particular, if (1) and (2) are true, then (3) must be true too. Notice that the validity of this pattern only depends on the meaning of the connective $\rightarrow$ and not on the meaning of the formulas P_1, P_2. We call the concrete argument about snow and being cold *correct*, because the underlying reasoning pattern is valid.

But, for instance, the validity of the reasoning pattern

For all x, if x is a person, then x is mortal	$\forall x[P(x) \rightarrow M(x)]$
Socrates is a person	$P(c)$
Therefore: Socrates is mortal	$M(c)$

not only depends on the meaning of the connective $\rightarrow$, but also on the meaning of the universal quantifier $\forall$ (for all). Notice that P_1, P_2 above stand for propositions, while $P(x), M(x)$ stand for the predicates 'x is a person' and 'x is mortal' respectively. In predicate logic, to be treated in Chapter 4, we study reasoning patterns of which the validity also depends on the meaning of the quantifiers $\forall$ (for all) and $\exists$ (for at least one).

The study of propositional logic was initiated by the Stoics (see Subsection 2.10.2), some 300 years before Aristotle developed his theory of the syllogisms (see Subsection 4.7.4).

Let us start by considering some examples of propositions about numbers, some of which are true and some of which are false. We give their translation into the language of propositional logic, and their translation into the language of predicate logic.

Proposition	*prop. formula*	*pred. formula*
1. All numbers are positive (≥ 0)	P_1	$\forall x[P(x)]$
2. All numbers are negative (≤ 0)	P_2	$\forall x[N(x)]$
3. All numbers are positive or negative	P_3	$\forall x[P(x) \vee N(x)]$

Here $\forall$ is the *universal quantifier* expressing 'for all', $P(x)$ stands for the predicate 'x is positive', $N(x)$ for the predicate 'x is negative' and $\vee$ stands for the connective 'or'. It is important to notice that the propositional translation of sentence 3 cannot be rendered by $P_1 \vee P_2$, because this formula expresses the proposition 'all numbers are positive or all numbers are negative' which happens to be false, while sentence 3 is true. Also notice that in $P_1 \vee P_2$ the connective $\vee$ stands between two propositions, while in $\forall x[P(x) \vee N(x)]$ the connective $\vee$ stands between two predicates.

Proposition	*prop. formula*	*pred. formula*
4. There is at least one even number	P_4	$\exists x[E(x)]$
5. There is at least one odd number	P_5	$\exists x[O(x)]$
6. There is a number that is both even and odd	P_6	$\exists x[E(x) \wedge O(x)]$

Here $\exists$ is the *existential quantifier* expressing 'there is at least one', $E(x)$ stands for the predicate 'x is even', $O(x)$ for the predicate 'x is odd' and $\wedge$ stands for the connective 'and'. It is important to notice that the propositional translation of sentence 6 cannot be rendered by $P_4 \wedge P_5$, because this formula expresses the proposition 'there is at least one even number and there is at least one odd number' which happens to be true, while sentence 6 is false. Also notice that in $P_4 \wedge P_5$ the connective $\wedge$ stands between two propositions, while in $\exists x[E(x) \wedge O(x)]$ the connective $\wedge$ stands between two predicates.

Proposition	*prop. formula*	*pred. formula*
7. There is a number x such that $x > 0$	P_7	$\exists x[x > 0]$
8. There is a number x such that not $x > 0$	P_8	$\exists x[\neg(x > 0)]$

$x > 0$ is not a proposition, but a predicate, while $5 > 0$, for instance, is a proposition. Similarly, 'not $x > 0$' is not a proposition, but the negation of a predicate, while 'not $5 > 0$' is a proposition. It is important to notice that proposition 8 is not the negation of proposition 7; the negation of 7 is 'there is no number x such that $x > 0$', $\neg\exists x[x > 0]$, which is equivalent to 'for all numbers x, not $x > 0$'. This latter proposition is false, while proposition 8 is true. In the negation of sentence 7, the negation stands in front of the existential quantifier, while in sentence 8 the negation stands in front of the predicate $x > 0$.

Proposition	*prop. formula*	*pred. formula*
9. All persons have a mother	P_9	$\forall x \exists y[M(x,y)]$
10. There is one mother of all persons	P_{10}	$\exists y \forall x[M(x,y)]$

$\forall x \exists y[M(x,y)]$ says: for every person x there is a person y such that x stands in the child-mother relation $M(x,y)$ with y. But by changing the order of the quantifiers one obtains $\exists y \forall x[M(x,y)]$ which says: there is at least one person y such that for all persons x, x stands in the child-mother relation $M(x,y)$ with y. Notice that sentence 9 is true, while sentence 10 is false.

Proposition	*prop. formula*	*pred. formula*
11. For every number there is a larger one	P_{11}	$\forall x \exists y[x < y]$
12. There is a largest number	P_{12}	$\exists y \forall x[x < y]$

$\forall x \exists y[x < y]$ says: for every number x there is a number y such that x is smaller than y. But changing the order of the quantifiers one obtains $\exists y \forall x[x < y]$ which says: there is a number y such that for all numbers x, x is smaller than y. Notice that sentence 11 is true, while sentence 12 is false. So, the order of the quantifiers $\forall$ and $\exists$ does matter!

Let us have a closer look at proposition 9: 'all persons have a mother', or equivalently:

For every person x there is some person y such that y is the mother of x.

I

II

III

I, 'y is the mother of x', does not express a proposition, but a binary predicate or relation; neither does 'Mary is the mother of x', which expresses a unary predicate. However, 'Mary is the mother of John' does express a proposition.
II, 'there is some person y such that y is the mother of x', does not express a proposition, but a unary predicate, which may become more clear if we formulate II as follows: someone is the mother of x or, equivalently, x has a mother. However, 'someone is the mother of John' does express a proposition.
III does express the proposition 'every person has a mother'. Note that all variables x, y occurring in III also occur in the context 'for every' or 'there is'.

In *propositional logic* one ignores the internal subject-predicate structure of the atomic propositions. The atomic propositions can have the form 'for all x, x has a certain property P', like the propositions 1 up to 3 inclusive, or the form 'there is at least one x such that x has the property P', like the propositions 4 up to 8 inclusive, or the form 'for every x there is a y such that x is in relation $R(x,y)$ to y', like proposition 9 and 11, and so on. In the *propositional calculus* we restrict ourselves to arguments like Modus Ponens and the arguments a), b) and c) in Chapter 1, the correctness of which only depends on how the different propositions are composed of more elementary ones by means of operations like 'iff', 'if ..., then ...', 'and', 'or' and 'not'. In the propositional calculus the internal subject-predicate structure of the elementary propositions is not taken into consideration. However, the argument above about Socrates makes it clear that the correctness of an argument may also depend on this subject-predicate structure. Therefore, the propositional calculus has to be extended to the predicate calculus, which is treated in Chapter 4.

Below we list the symbols we are using for the propositional operations, mentioning their name and alternative symbols which may be used in the literature.

name	*symbol*	*alternatives*	*meaning*
equivalence	$\rightleftarrows$	$\leftrightarrow, \sim, \equiv$	(is) equivalent (to); if and only if; iff
(material) implication	$\rightarrow$	$\supset$	if ..., then ...; implies
conjunction	$\wedge$	&	and
disjunction	$\vee$		or; and/or
negation	$\neg$	-	not

Instead of the atomic propositions considered above, being about numbers, and the propositions that can be built from them by the propositional connectives, we can of course consider different atomic propositions, for instance of geometry, physics or of some other sharply circumscribed part of natural language, together with the composite propositions that can be built from them. So, in order to retain flexibility for the applications, we shall simply assume, throughout this chapter, that we are dealing with an object language in which there is a class of (declarative) sentences consisting of certain building blocks

$$P_1,\ P_2,\ P_3, \ldots$$

called *atomic formulas*, from which *composite formulas* can be built by means of the propositional connectives. By a *formula* we mean either an atomic or a composite formula. So, throughout this chapter our object language will be the following symbolic language:

	symbols	*names*
Alphabet:	$P_1,\ P_2,\ P_3, \ldots$	atomic formulas or propositional variables
	$\rightleftarrows,\ \rightarrow,\ \wedge,\ \vee,\ \neg$	connectives
	(,)	parentheses

Definition 2.1 (Formulas).

1. Each atomic formula is a formula. In other words, if P is an atomic formula, then P is a formula.

2. If each of A and B is a given formula (i.e., either an atomic formula or a composite formula already constructed), then $(A \rightleftarrows B)$, $(A \to B)$, $(A \wedge B)$ and $(A \vee B)$ are (composite) formulas.
3. If A is a given formula, then $(\neg A)$ is a (composite) formula.

This language is the *formal language of propositional logic*. It consists of the formulas built from the given alphabet according to Definition 2.1.

P_1, P_2, P_3, ... are symbols to be interpreted as atomic propositions from arithmetic, geometry, physics, any other science or daily life. The first four connectives are binary connectives, the last one is unary. The connectives are symbols whose meanings are the respective propositional operations; in Section 2.2 we will fix and stylize these meanings by truth tables. The parentheses are punctuation marks. In $A \to B$ we call A the *antecedent* and B the *succedent*.

Example 2.1. Here are some examples of formulas:

$$P_1,\ P_2,\ P_3,\ P_4$$
$$(\neg P_2),\ (\neg P_3)$$
$$(P_1 \vee (\neg P_2)),\ ((\neg P_3) \to P_4)$$
$$((P_1 \vee (\neg P_2)) \wedge ((\neg P_3) \to P_4)).$$

Notice that the number of left parentheses must be equal to the number of right parentheses.

if, only if and **iff**:
'B if A' is translated by $A \to B$, which may also be read as 'if A, then B'.
'B only if A' is translated by $B \to A$, and 'B if and only if (iff) A' is translated by $(A \to B) \wedge (B \to A)$, or, equivalently, by $A \rightleftarrows B$.

Convention: When we want to state something about arbitrary natural number, the letters n, m are used to stand for any of the natural numbers 0, 1, 2, 3, ... For instance, when we state that for all natural numbers n, m: $n+m = m+n$. Similarly, the letters P, Q and R are used to stand for any atomic formulas $P_1, P_2, P_3, \ldots$ and the letters $A, B, C, A_1, A_2, \ldots, B_1, B_2, \ldots, C_1, C_2, \ldots$ are used to stand for any formulas, not necessarily atomic. For instance, the letters A and B may stand for any of the formulas in Example 2.1. Distinct such letters need *not* represent distinct formulas, in contrast to $P_1, P_2, P_3, \ldots$ which are distinct atomic formulas.

Parentheses in formulas are essential: they indicate which parts belong together. Leaving them out may cause ambiguity. For instance, $A \wedge B \to C$ might mean:

- $(A \wedge B) \to C$, which is an implicational formula with $A \wedge B$ as antecedent, and
- $A \wedge (B \to C)$, which is a conjunction of the formulas A and $B \to C$.

'If John wins the lottery and is healthy, then he will go to the Bahamas' is a proposition of the first form, while 'John wins the lottery and if he is healthy, then he will go to the Bahamas' is a proposition of the second form. Only in the second proposition it is stated that John wins the lottery.

Convention We can minimize the need for parentheses by agreeing that we leave out the most outer parentheses in a formula and that in

$$\rightleftarrows,\ \rightarrow,\ \wedge,\ \vee,\ \neg$$

any connective has a higher rank than any connective to the right of it and a lower rank than any connective to the left of it.

According to this convention, $A \wedge B \rightarrow C$ should be read as $(A \wedge B) \rightarrow C$, because $\rightarrow$ has a higher rank than $\wedge$, and not as $A \wedge (B \rightarrow C)$, which has a different meaning. The formula $\neg A \vee B$ should be read as $(\neg A) \vee B$, because by convention $\vee$ has a higher rank than $\neg$, and not as $\neg(A \vee B)$, which means quite something else. And $C \rightleftarrows A \wedge B \rightarrow C$ should be read as $C \rightleftarrows ((A \wedge B) \rightarrow C)$.

It is interesting to notice that the build-up of formulas is very similar to the build-up of natural numbers. Formulas are generated by starting with atomic formulas P_1, P_2, P_3, ... and successively passing from one or two formulas already generated before to another formula by means of the connectives. Natural numbers are generated by starting with one initial object 0 and successively passing from a natural number n already generated before to another natural number $n+1$ or n' (the *successor* of n).

Since natural numbers are built up from 0 by repeated application of the successor operation, the theorem of *mathematical induction* follows immediately from the definition of natural numbers:

Theorem 2.1 (Mathematical induction). *Let Φ be a property of natural numbers such that*

1. *(induction basis:) 0 has property Φ, and*
2. *property Φ is preserved when going from a natural number n to its successor n', i.e., for all natural numbers n, if n has property Φ (induction hypothesis), then also n' has property Φ.*

Then all natural numbers have property Φ.

Using mathematical induction, one can prove, for instance, that for all natural numbers n, $1 + 2 + \ldots + n = \frac{1}{2}n(n+1)$. See Exercise 2.5.

The *induction principle* for formulas is similar to mathematical induction for natural numbers. Since (propositional) formulas are built up from atomic formulas $P_1, P_2, P_3, \ldots$ by successive applications of connectives to formulas already generated before, the following theorem, called the *induction principle* (for propositional formulas), follows immediately from the definition of formulas.

Theorem 2.2 (Induction principle). *Let Φ be a property of formulas, satisfying*

1. *(induction basis:) every atomic formula has property Φ and*
2. *property Φ is preserved in building more complex formulas by means of the connectives, i.e., if A and B have property Φ (induction hypothesis), then $(A \rightleftarrows B)$, $(A \rightarrow B)$, $(A \wedge B)$. $(A \vee B)$ and $(\neg A)$ also have property Φ.*

Then every formula (of the propositional calculus) has property Φ.

Using Theorem 2.2 one can prove, for instance, that every formula contains as many left parentheses as right parentheses (see Exercise 2.6.) Another application is Theorem 2.18 which says that every formula can be written in normal form.

Notice that we have introduced a logical (propositional) language such that English sentences may be translated into this logical language and conversely one may translate the logical formulas into the corresponding English sentences. What holds for English sentences of course also holds for German, French, Spanish and all other sentences. With this in mind one might build for each natural language a machine that translates the sentences of the language in question into logical formulas and back. By combining these machines with logic as the intermediate language, one obtains an automatic translation of, for instance, English to, for instance, German: automatically translate the English sentences into logical formulas and next automatically translate the resulting logical formulas into German sentences. This was roughly the Rosetta translation project of the European Union.

Exercise 2.1. Let P_1 stand for 'John works hard',
P_2 for 'John is going to school', and
P_3 for 'John is wise'.
Translate the following sentences into the language of propositional logic, using the least possible number of parentheses.

i) If John works hard and is going to school, then John is not wise.
ii) John works hard and if John is going to school, then he is not wise.
iii) John works hard, or if John is going to school, then he is wise.
iv) If John is going to school or works hard, then John is wise.
v) If John works hard, then John is not wise, at least if he is going to school.

Exercise 2.2. Translate the following formulas into English sentences, reading P_1, P_2 and P_3 as indicated in exercise 2.1.

i) $(P_1 \to P_2) \to \neg P_3$
ii) $\neg P_1 \vee P_3$
iii) $\neg(P_1 \vee P_3)$
iv) $\neg P_2 \wedge P_3$
v) $\neg(P_2 \wedge P_3)$

Exercise 2.3. Translate the following propositions into propositional formulas and into predicate formulas:
1. Every gnome has a beard.
2. All gnomes have no beard.
3. Not every gnome has a beard.

Exercise 2.4. Which of the following expressions are formulas (of the language of propositional calculus)? P_1, P, $\neg P_8$, $\neg Q$, $P_1 \wedge \neg P_8$, $P \wedge \neg Q$, A, B, $A \wedge \neg B$, $(P_1 \wedge P_2) \to \neg P_3$, $(P_1 \wedge P_2) \to Q$, $(P_1 \wedge P_2) \to B$, $A \wedge B \to C$.

Exercise 2.5. Use mathematical induction (Theorem 2.1) to prove that for all natural numbers n, $1 + 2 + \ldots + n = \frac{1}{2}n(n+1)$.

Exercise 2.6. Use the induction principle (Theorem 2.2) to show that every formula of propositional logic contains as many left parentheses '(' as right parentheses ')'.

2.2 Semantics; Truth Tables

In the first section of this chapter a logical (propositional) language was introduced in which we can translate the premisses and the conclusion of an argument, resulting in a reasoning pattern. We have indicated the meaning of the atomic formulas: atomic propositions which are either true or false. And we have indicated the meaning of the propositional connectives $\rightleftarrows$, $\to$, $\wedge$, $\vee$, and $\neg$: 'if and only if', 'if ..., then ...', 'and', 'or', and 'not', respectively.

In this section the meaning of the atomic formulas and the propositional connectives is made more precise, where we restrict ourselves (in this chapter) to classical logic. Owing in part to different analyses of implication, the heart of logic, there are different systems of logic: classical logic, intuitionistic logic, relevance logic and so on. Although we will treat the latter logic systems in other chapters, in this chapter we shall concern ourselves primarily with *classical logic*, because it is the simplest and most commonly used system of logic. In classical logic we assume that each proposition is either *true*, indicated by 1, or *false* indicated by 0. We do not, however, suppose that one always *knows* whether a particular proposition is true or false.

To start with, the atomic formulas P_1, P_2, P_3, ... stand for (or are interpreted as) atomic propositions, such as 'John is ill', the 'weather is nice', etc. These atomic propositions may be true, indicated by 1, or false, indicated by 0. We standardize this in the so-called *truth table* of the atomic formulas $P_1, P_2, P_3, \ldots$. So, by definition the truth table of an atomic formula P, where P stands for any of the atomic formulas P_1, P_2, P_3, ..., is the following one:

P
1
0

For two atomic formulas P and Q there are four different assignments of the values 1 (true) and 0 (false), schematically rendered as follows:

P	Q
1	1
1	0
0	1
0	0

In the first line the atomic formulas P and Q are both interpreted as true atomic propositions, in the fourth line both as false atomic propositions.

For three atomic formulas P, Q and R there are eight different assignments of the values 1 and 0. Notice that the number of different assignments of the values 1 and 0 to P, Q and R is two times as many as for the two atomic formulas P and Q, since for each of the four different assignments of the values 1 and 0 to P and Q, one may assign a 1 or a 0 to R: More generally:

Lemma 2.1. *For n atomic formulas $P_1, \ldots, P_n$, $n = 1, 2, \ldots$, there are 2^n different assignments of the values 1 and 0.*

If a formula A, for instance $A = P_1 \to (P_2 \to P_3)$, has been built from three atomic formulas, there are $2^3 = 8$ different assignments of the values 1 and 0 to the atomic formulas P_1, P_2 and P_3. But the formula A itself can have at most two different values: 1 and 0.

Next a precise meaning has to be given to the propositional connectives. This is done in the so-called *truth tables for the propositional connectives*, where it is specified how the truth value of the composite formulas $A \rightleftarrows B$, $A \to B$, $A \wedge B$, $A \vee B$ and $\neg A$ is completely determined by the truth values of the components A and B.

Two different formulas A and B can have at most four different values of truth (1) and falsity (0), represented by the four rows in the table below. Each column in the table below indicates how the truth or falsity of the composite formula heading that column depends on the truth values of its immediate components A and B.

A	B	$A \rightleftarrows B$	$A \to B$	$A \wedge B$	$A \vee B$	A	$\neg A$
1	1	1	1	1	1	1	0
1	0	0	0	0	1	0	1
0	1	0	1	0	1		
0	0	1	1	0	0		

Thus $A \rightleftarrows B$ is true exactly when A and B have the same truth value; hence, the reading 'equivalent', i.e., 'equal valued', for $\rightleftarrows$.
$A \to B$ is false exactly when A is true and B is false.
$A \wedge B$ is true exactly when A and B are both true.
$A \vee B$ is false exactly when both A and B are false.
And $\neg A$ is true exactly when A is false.

The truth tables for the propositional connectives may also be presented in the following way:

		$\rightleftarrows$	$\to$	$\wedge$	$\vee$		$\neg$
1	1	$1 \rightleftarrows 1 = 1$	$1 \to 1 = 1$	$1 \wedge 1 = 1$	$1 \vee 1 = 1$	1	$\neg 1 = 0$
1	0	$1 \rightleftarrows 0 = 0$	$1 \to 0 = 0$	$1 \wedge 0 = 0$	$1 \vee 0 = 1$	0	$\neg 0 = 1$
0	1	$0 \rightleftarrows 1 = 0$	$0 \to 1 = 1$	$0 \wedge 1 = 0$	$0 \vee 1 = 1$		
0	0	$0 \rightleftarrows 0 = 1$	$0 \to 0 = 1$	$0 \wedge 0 = 0$	$0 \vee 0 = 0$		

The truth tables for $\rightleftarrows$, $\wedge$, $\vee$ and $\neg$ are self evident and give little or no reason for discussion. However, the table for $\to$ was already disputed by the Stoics, see Subsection 2.10.2. Nevertheless, it is the only one of the 16 possible columns of length 4 consisting of 1's and 0's which is tenable; any other proposal can easily be rejected as unreasonable.

First, let us notice that the propositional connectives $\rightleftarrows$, $\to$, $\wedge$, $\vee$ and $\neg$ as defined in the truth tables are *truthfunctional*, i.e., the truth values of $A \rightleftarrows B$, $A \to B$, $A \wedge B$, $A \vee B$ and $\neg A$ are completely determined by the truth values of its components A and B. This is not always the case for the connective 'if ..., then ...' from daily language, as may be illustrated by the following two sentences:

1. If I would have jumped out of the window on the 10th floor, then I would have been injured.

2. If I would have jumped out of the window on the 10th floor, then I would have changed into a bird.
Although in both sentences the components have the same truth value 0 (I have not jumped out of the window, I have not been injured and I have not changed into a bird) the first sentence is held to be true, while the second sentence is held to be false. In other words, in sentence 1, the combination 'if 0, then 0' gives a 1, while in sentence 2 the same combination 'if 0, then 0' gives a 0. So, the 'if ..., then ...' from daily language is not truthfunctional. Consequently, the $\rightarrow$ may be different from the 'if ..., then ...' from daily language.

Neverthelesse, in daily life the 'if ..., then ...' is frequently, although not always, used precisely as described in the truth table of $\rightarrow$. We may illustrate this with the following example:

For all integers n and m, if $n = m$, then $n^2 = m^2$.

Why is this proposition true? Simply because it is impossible that for some integers n and m the proposition $n = m$ has truth value 1, while the proposition $n^2 = m^2$ has truth value 0. In other words, the combination 1 for $n = m$ and 0 for $n^2 = m^2$ does not occur. Only the combinations 1 - 1, 0 - 1 and 0 - 0 may occur and these give the value 1, just as in the truth table of $\rightarrow$:

	$n = m$	$n^2 = m^2$	if $n = m$, then $n^2 = m^2$
$n = 2,\ m = 2$	1	1	1
$n = 2,\ m = -2$	0	1	1
$n = 2,\ m = 3$	0	0	1

From the table for $\rightarrow$ one sees that $A \rightarrow B$ is true (has value 1; is 1) if and only if A is false ($\neg A$ is true) or B is true (has value 1); in other words, it is easy to check that $A \rightarrow B$ and $\neg A \vee B$ have the same truth table. The truth table of $A \rightarrow B$ is also the same as the one of $\neg(A \wedge \neg B)$, which corresponds with our intuitions:

A	B	$\neg A$	$\neg A \vee B$	$\neg B$	$A \wedge \neg B$	$\neg(A \wedge \neg B)$
1	1	$\neg 1 = 0$	$0 \vee 1 = 1$	$\neg 1 = 0$	$1 \wedge 0 = 0$	$\neg 0 = 1$
1	0	$\neg 1 = 0$	$0 \vee 0 = 0$	$\neg 0 = 1$	$1 \wedge 1 = 1$	$\neg 1 = 0$
0	1	$\neg 0 = 1$	$1 \vee 1 = 1$	$\neg 1 = 0$	$0 \wedge 0 = 0$	$\neg 0 = 1$
0	0	$\neg 0 = 1$	$1 \vee 0 = 1$	$\neg 0 = 1$	$0 \wedge 1 = 0$	$\neg 0 = 1$

Warning One frequently is inclined to read $A \rightarrow B$ as: A and hence B. But this is wrong! If I assert $A \rightarrow B$, I do not assert A, neither B. Consider, for instance, the sentence: if I win the lottery, then I will give you a Cadillac. This does not mean that I win the lottery and hence will give you a Cadillac.

Why is $A \rightarrow B$ true (1) in case A is false (0)? Consider the following example. Suppose I am determined never to play in a lottery; in this case I can truthfully state: If I win the lottery, then I will give you a Cadillac. Assuming I never play in a lottery, this is an empty statement, without content, and hence this statement cannot be false.

And why is $A \to B$ true (1) if B is true (1)? Suppose B stands for 'I give you a Cadillac' and suppose this is true (1). Then the sentence 'if I win the lottery, then I will give you a Cadillac' is certainly true (1) too.

The reader should also verify that the truth table for $A \rightleftarrows B$ is the same as the one of $(A \to B) \wedge (B \to A)$, which also corresponds with our intuition:

A	B	$A \rightleftarrows B$	$A \to B$	$B \to A$	$(A \to B) \wedge (B \to A)$
1	1	1	1	1	1
1	0	0	0	1	0
0	1	0	1	0	0
0	0	1	1	1	1

If one constructs the truth tables for $A \wedge B$ and for $B \wedge A$, one will find that these two truth tables are the same:

A	B	$A \wedge B$	$B \wedge A$
1	1	$1 \wedge 1 = 1$	$1 \wedge 1 = 1$
1	0	$1 \wedge 0 = 0$	$0 \wedge 1 = 0$
0	1	$0 \wedge 1 = 0$	$1 \wedge 0 = 0$
0	0	$0 \wedge 0 = 0$	$0 \wedge 0 = 0$

However, a sentence like 'Ann had a baby and got married' will leave another impression than the sentence 'Ann got married and had a baby'. In this example the order of the two atomic propositions suggests a temporal or causal succession. Also in the sentence 'John fell into the water and drowned' one cannot easily change the order of the atomic components. These examples show that the connectives from daily language may have shades of meaning which are lost in their translation to the corresponding propositional connectives. Notice that the expression 'A but B' has nuances of meaning not possessed by 'A and B' and lost in the translation $A \wedge B$: 'I love you and I love your sister almost as well' will leave another impression than 'I love you but I love your sister almost as well'.

In daily life, the connective 'or' is sometimes used in an exclusive way. For instance, when the dinner menu says 'tea or coffee included', we do not expect to get both. But in 'books can be delivered at school or at church' the connective 'or' is used in an inclusive way: we may deliver books at school and/or at church. Notice that the symbol $\vee$, coming from the Latin 'vel', corresponds with the inclusive 'or' and that $A \vee B$ has the same truth table as $B \vee A$.

Analysing the use of the propositional operations 'iff', 'if ..., then ...', 'and', 'or', and 'not' in arithmetic, calculus and more generally in mathematics, it turns out that these operations are used precisely as described in the truth tables of $\rightleftarrows$, $\to$, $\wedge$, $\vee$ and $\neg$ respectively. This should make it clear that our propositional connectives and *material implication* $A \to B$ in particular are useful and natural forms of expression. In natural language the propositional operations are frequently, but not always, used as described in the truth tables above.

No disagreement exists that 'if A, then B' is false if A is true and B is false. Problems arise with the claim that 'if A, then B' is false *only* if A is true and B is false, and is true in all other cases. 'If these three chairs cost 6 dollars (A), then one chair costs 2 dollars (B)' is true, because it is impossible that A is true and B is false, due to the causal relation between A and B; in this example both A and B are supposed to be false. Problems arise if there is no connection of ideas between A and B, like in 'if I would have jumped out of the window, then I would have changed into a bird', which is true under our table. $A \to B$ is called a *conditional* or a *material implication*; the latter name because the truth of 'if A, then B' in general depends on matters of empirical fact.

Example 2.2. Let us illustrate the repeated use of the truth tables by computing the one for $P_1 \to (P_2 \to P_3)$ and the one for $(P_1 \wedge P_2) \to P_3$:

P_1	P_2	P_3	$P_2 \to P_3$	$P_1 \to (P_2 \to P_3)$	$P_1 \wedge P_2$	$(P_1 \wedge P_2) \to P_3$
1	1	1	$1 \to 1 = 1$	$1 \to 1 = 1$	$1 \wedge 1 = 1$	$1 \to 1 = 1$
1	1	0	$1 \to 0 = 0$	$1 \to 0 = 0$	$1 \wedge 1 = 1$	$1 \to 0 = 0$
1	0	1	$0 \to 1 = 1$	$1 \to 1 = 1$	$1 \wedge 0 = 0$	$0 \to 1 = 1$
1	0	0	$0 \to 0 = 1$	$1 \to 1 = 1$	$1 \wedge 0 = 0$	$0 \to 0 = 1$
0	1	1	$1 \to 1 = 1$	$0 \to 1 = 1$	$0 \wedge 1 = 0$	$0 \to 1 = 1$
0	1	0	$1 \to 0 = 0$	$0 \to 0 = 1$	$0 \wedge 1 = 0$	$0 \to 0 = 1$
0	0	1	$0 \to 1 = 1$	$0 \to 1 = 1$	$0 \wedge 0 = 0$	$0 \to 1 = 1$
0	0	0	$0 \to 0 = 1$	$0 \to 1 = 1$	$0 \wedge 0 = 0$	$0 \to 0 = 1$

Notice that $P_1 \to (P_2 \to P_3)$ has the same truth table as $(P_1 \wedge P_2) \to P_3$, which corresponds with our intuition: $P_1 \to (P_2 \to P_3)$ is read as 'if P_1, then (if - in addition - P_2, then P_3), which is equivalent to 'if P_1 and P_2, then P_3'.

2.2.1 Validity

Atomic formulas have (by definition) two truth values, 1 and 0. However, it is easy to see that some composite formulas have only one truth value. For instance, the formula $P_1 \to P_1$ can only have the truth value 1, no matter what the truth value of P_1 is. And the formula $P_1 \wedge \neg P_1$ can only have the truth value 0, no matter what the truth value of P_1 is:

P_1	$P_1 \to P_1$	$\neg P_1$	$P_1 \wedge \neg P_1$
1	$1 \to 1 = 1$	$\neg 1 = 0$	$1 \wedge 0 = 0$
0	$0 \to 0 = 1$	$\neg 0 = 1$	$0 \wedge 1 = 0$

Other formulas with only the truth value 1 are $P_1 \vee \neg P_1$, $P_1 \wedge P_2 \to P_1$, $P_1 \to (P_2 \to P_1)$ and $P_1 \to P_1 \vee P_2$. These formulas are called *always true* or *valid.* Wittgenstein (1921) called these formulas *tautologies*.

Definition 2.2 (Valid; Consistent; Contingent). Let A be a formula.
A is *always true* or *valid* := the truth table of A – entered from the atomic formulas from which A has been built – contains only 1's. **Notation**: $\models A$.
A is *consistent* or *satisfiable* := the truth table of A contains at least one 1; that is, the formula A may be true.

A is *contingent* := the truth table of A contains at least one 1 and at least one 0; that is, A may be true and it may be false.
A is *inconsistent* or *always false* or *contradictory* := the truth table of A contains only 0's; that is, A cannot be true, in other words, is always false.

Notice that a valid formula is consistent, but not contingent; that a contingent formula is by definition also consistent; and that an inconsistent formula is by definition not contingent.

So, for instance, the formula $P_1 \to P_1$ is valid and hence also consistent, the formula $P_1 \to P_2$ is contingent and consistent, but not valid, and the formula $P_1 \wedge \neg P_1$ is inconsistent or always false.

On the one hand, valid formulas are uninteresting because they give no information. On the other hand, since valid formulas are always true regardless of the truth or falsity of their atomic components, they may be used in reasoning as may be illustrated by the following example.

Example 2.3. Consider the following argument:

John is lazy [L].	L
If John is ill [I] or lazy, he stays at home [H].	$\underline{I \vee L \to H}$
Therefore: John stays at home.	H

In this valid reasoning pattern we use silently that $\models L \to I \vee L$. The argument might be simulated as follows:

$$\dfrac{\dfrac{L \quad L \to I \vee L}{I \vee L} \quad I \vee L \to H}{H}$$

Note that there are infinitely many valid formulas. Although it is not exhaustive (for instance, $P \vee \neg P$ does not occur in it), the following list enumerates infinitely many valid formulas.

$P \to P$
$P \to (P \to P)$
$P \to (P \to (P \to P))$
$\vdots$

Warning: While the symbol A stands for any formula, like $P_1 \to P_2$, $P_2 \wedge \neg P_3$, etc., the expression $\models A$ is not a formula, but a statement *about* the formula A, namely, that the truth table of A contains only 1's. The symbol $\models$ does not occur in the logical alphabet, and '$\models A$' is shorthand for 'A is valid' or 'A is always true', which clearly is not a logical formula. In other words, the symbol A indicates a formula from the logical language, our object language, while the expression $\models A$ belongs to the meta-language, in which we make statements *about* formulas of the object language.

Notation: If a particular formula A is not valid, this is frequently written by $\not\models A$ instead of 'not $\models A$'. For instance: $\models P_1 \to P_1 \vee P_2$, but $\not\models P_1 \to P_1 \wedge P_2$.

Definition 2.3 (Interpretation; Model). Let A be a formula built from the atomic formulas $P_1, \ldots, P_n$. An *interpretation* i of A assigns a value 1 or 0 to all the atomic components of A; so, an interpretation i of A corresponds with a line in the truth table for A and interprets each atomic formula in A as either a true or a false proposition. Interpretation i of A is a *model* of A := i assigns to A the value 1, in other words, $i(A) = 1$. In this terminology the definition of 'A is valid' can be reformulated as follows: every interpretation i of A is a model of A.

Example 2.4. Thus, if A has been built from only two atomic formulas P and Q, then there are four different interpretations of A: i_1, i_2, i_3, i_4.

	P	Q	$P \to Q$	
i_1	1	1	1	$i_1(P) = 1, i_1(Q) = 1, i_1(P \to Q) = 1$
i_2	1	0	0	$i_2(P) = 1, i_2(Q) = 0, i_2(P \to Q) = 0$
i_3	0	1	1	$i_3(P) = 0, i_3(Q) = 1, i_3(P \to Q) = 1$
i_4	0	0	1	$i_4(P) = 0, i_4(Q) = 0, i_4(P \to Q) = 1$

For instance, i_1, i_3 and i_4 are a model of $P \to Q$, but i_2 is not a model of $P \to Q$.

Definition 2.4. Let Γ be a (possibly infinite) set of formulas and i an interpretation, assigning the values 0 or 1 to all the atomic components of the formulas in Γ.
i a *model* of Γ := i is a model of all formulas in Γ, i.e., i makes all formulas in Γ true.
Γ is *satisfiable* := there is at least one assignment i which is a model of Γ.

Example 2.5. If Γ consists of $P_1 \to P_2$ and $P_1 \vee P_2$, then i_1 and i_3 are models of Γ.

	P_1	P_2	$P_1 \to P_2$	$P_1 \vee P_2$
i_1	1	1	1	1
i_2	1	0	0	1
i_3	0	1	1	1
i_4	0	0	1	0

Theorem 2.3 (Compactness theorem). * *Let Γ be a (possibly infinite) set of formulas such that every finite subset of Γ has a model. Then Γ has a model.*

Proof. Let Γ be a (possibly infinite) set of formulas such that every finite subset of Γ has a model. We will define an interpretation i of the atomic propositional formulas $P_1, P_2, P_3, \ldots$ such that for every natural number n, $\Phi(n)$, where $\Phi(n)$:= every finite subset of Γ has a model in which $P_1, P_2, \ldots, P_n$ take the values $i(P_1), i(P_2), \ldots, i(P_n)$.

Once having shown this, it follows that $i(A) = 1$ for every formula A in Γ. For given a formula A in Γ, take n so large that all atomic formulas occurring in A are among $P_1, \ldots, P_n$. Since $\{A\}$ is a finite subset of Γ and because of $\Phi(n)$, A has a model in which $P_1, \ldots, P_n$ take the values $i(P_1), \ldots, i(P_n)$. So, $i(A) = 1$.

Let $i(P_1) = 0$ and suppose $\Phi(1)$ does not hold. That is, there is a finite subset Γ' of Γ which has no model in which P_1 takes the value $i(P_1) = 0$. Then we define $i(P_1) = 1$ and show that $\Phi(1)$, i.e., every finite subset of Γ has a model in which P_1 takes the value $i(P_1) = 1$. For let Δ be a finite subset of Γ. Then $\Delta \cup \Gamma'$ is a finite subset of Γ and hence has a model i. Since i is a model of Γ', $i(P_1) = 1$.

Suppose we have defined $i(P_1), \ldots, i(P_n)$ such that $\Phi(n)$. Then we can extend the definition of i to P_{n+1} such that $\Phi(n+1)$. For suppose that $\Phi(n+1)$ does not hold if $i(P_{n+1}) = 0$. That is, there is a finite subset Γ' of Γ which has no model in which $P_1, \ldots, P_n, P_{n+1}$ take the values $i(P_1), \ldots, i(P_n), 0$. Then we define $i(P_{n+1}) = 1$ and show that $\Phi(n+1)$, i.e., every finite subset of Γ has a model in which $P_1, \ldots, P_n, P_{n+1}$ take the values $i(P_1), \ldots, i(P_n), 1$. For let Δ be a finite subset of Γ. Then $\Delta \cup \Gamma'$ is a finite subset of Γ and hence, by the induction hypothesis, $\Delta \cup \Gamma'$ has a model in which $P_1, \ldots, P_n$ take the values $i(P_1), \ldots, i(P_n)$. Since i is a model of Γ', $i(P_{n+1}) = 1$. □

For applications of the compactness theorem in mathematics see Exercises 2.16, 2.17 and 2.18.

Exercise 2.7. Show that the formulas in the pairs below have the same truth table:

a) $\neg(A \wedge B)$ and $\neg A \vee \neg B$.
b) $\neg(A \vee B)$ and $\neg A \wedge \neg B$.
c) $\neg A \vee B$ and $A \to B$.
d) $\neg(A \to B)$ and $A \wedge \neg B$.
e) $A \to B$ and $\neg B \to \neg A$.
f) $A \to B$ and $\neg(A \wedge \neg B)$.

Exercise 2.8. Compute and compare the truth tables for:
a) $P_1 \wedge P_2 \to \neg P_3$ and $P_1 \wedge (P_2 \to \neg P_3)$ (see Exercise 2.1).
b) $P_1 \vee (P_2 \to P_3)$ and $P_1 \vee P_2 \to P_3$ (see Exercise 2.1).
c) $P_1 \to (P_2 \to \neg P_3)$ and $(P_1 \to P_2) \to \neg P_3$ (see Exercise 2.1 and 2.2).
d) $\neg P_1 \vee P_3$ and $\neg(P_1 \vee P_3)$ (see Exercise 2.2).
e) $\neg P_2 \wedge P_3$ and $\neg(P_2 \wedge P_3)$ (see Exercise 2.2).

Exercise 2.9. Prove that a) $(A \vee \neg A) \to B$ has the same truth table as B,
b) $(A \vee \neg A) \wedge B$ has the same truth table as B, and
c) $(A \wedge \neg A) \vee B$ has the same truth table as B.

Exercise 2.10. Prove that $A \vee B$, $(A \to B) \to B$ and $(B \to A) \to A$ all have the same truth table.

Exercise 2.11. Verify that the following formulas are valid by showing that it is impossible that at some line in the truth table they have the value 0.
a) $\neg\neg A \to A$ b) $(A \to B) \vee (B \to A)$ c) $(P \to Q) \to (\neg Q \to \neg P)$.

Exercise 2.12. Show that the following formulas are not valid by computing just one suitable line of the table: a) $P \vee Q \to P \wedge Q$ b) $(P \to Q) \to (Q \to P)$.

Exercise 2.13. Which of the following alternatives applies to the following formulas?
1. $P_1 \to \neg P_1$
2. $P_1 \rightleftarrows \neg P_1$
3. $P_1 \to P_1 \wedge P_2$
4. $P_1 \to P_1 \vee P_2$
5. $P_1 \to P_2$
6. $(P_1 \to P_2) \rightleftarrows (\neg P_1 \vee P_2)$
7. $\neg(P_1 \to P_2) \rightleftarrows (P_1 \wedge \neg P_2)$
8. $\neg(P_1 \wedge P_2) \rightleftarrows (\neg P_1 \vee \neg P_2)$
9. $\neg(P_1 \vee P_2) \rightleftarrows (\neg P_1 \wedge \neg P_2)$
10. $(\neg P_1 \vee P_2) \rightleftarrows (P_1 \to P_2)$

Alternative A: not satisfiable (inconsistent).
B: satisfiable (consistent), but not valid.
C: valid, and hence satisfiable.

Exercise 2.14. Show that each formula built by means of connectives from only one atomic formula P has the same truth table as either $P \wedge \neg P$, P, $\neg P$ or $P \to P$.

Exercise 2.15. Consider the following truth table for the *exclusive* 'or', $\underline{\vee}$.

A	B	$A \underline{\vee} B$
1	1	0
1	0	1
0	1	1
0	0	0

a) Verify that $A \underline{\vee} B$ has the same truth table as $(A \vee B) \wedge \neg(A \wedge B)$ and as $\neg(A \rightleftarrows B)$.
b) Verify that $(A \underline{\vee} B) \underline{\vee} C$ and $A \underline{\vee} (B \underline{\vee} C)$ have the same truth table and in particular that these formulas have the value 1 in the first line of the truth table (where A, B and C are 1). Note that this does not correspond with the intended meaning of 'A or B or C', if the 'or' is used exclusively.

Exercise 2.16. * (Kreisel-Krivine [18]) A group G is said to be *ordered* if there is a total ordering $<$ of G (see Chapter 3) such that $a \leq b$ implies $ac \leq bc$ and $ca \leq cb$ for all c in G. Show that a group G can be ordered if and only if every subgroup of G generated by a finite number of elements of G can be ordered.

Exercise 2.17. * (Kreisel-Krivine [18]) A graph (a non-reflexive symmetric relation) defined on a set V is said to be *k-chromatic*, where k is a positive integer, if there is a partition of V into k disjoint sets $V_1, \ldots, V_k$, such that two elements of V connected by the graph do not belong to the same V_i. Show that for a graph to be k-chromatic it is necessary and sufficient that every finite sub-graph be k-chromatic.

Exercise 2.18. * Suppose that each of a (possibly infinite) set of boys is acquainted with a finite set of girls. Under what conditions is it possible for each boy to marry one of his acquaintances? It is clearly necessary that every finite set of k boys be, collectively, acquainted with at least k girls. The *marriage theorem* says that this condition is also sufficient. More precisely, let B and G be sets (of Boys and Girls respectively) and let $R \subseteq B \times G$ be such that (i) for all $x \in B$, $R_{\{x\}}$ is finite, and (ii) for every finite subset $B' \subseteq B$, $R_{B'}$ has at least as many elements as B', where $R_{B'} := \{y \in G \mid \text{ for some } x \text{ in } B', R(x,y)\}$. Then there is an injection $f : B \to G$ such that for all $x \in B$ and $y \in G$, if $f(x) = y$, then $R(x, y)$. In **The Marriage Problem** (*American Journal of Mathematics*, Vol. 72, 1950, pp. 214-215) P. Halmos and H. Vaughan prove first the case in which the number of boys is finite. Using this result prove the marriage theorem for the case that B is infinite.

2.3 Semantics; Logical (Valid) Consequence

Consider the following concrete argument:

John is intelligent [I] or John is diligent [D].
If John is intelligent, then he will succeed [S].
If John is diligent, then he will succeed (too).
Therefore: John will succeed.

We may translate the propositions in this argument into formulas:

$$\frac{\begin{array}{c} I \vee D \\ I \to S \\ D \to S \end{array}}{S}$$

To help our memory, for convenience we have used the symbols I, D and S instead of P_1, P_2, P_3. Intuitively, this pattern of reasoning is valid: no matter what propositions I, D, S stand for, if all premisses are true, the conclusion must be true too; in other words, it is impossible that the premisses are all true and at the same time the conclusion false. Now we have given in Section 2.2 a precise meaning to the atomic formulas and to the connectives in terms of truth tables, we can make the notion of *valid* or *logical consequence* precise: in the truth table starting with I, D and S, at each line in which all of $I \vee D$, $I \to S$ and $D \to S$ have the value 1, also S must have the value 1; in other words: there is no line in the truth table starting with I, D, and S in which the premisses $I \vee D$, $I \to S$, $D \to S$ are all 1 and the conclusion S is 0.

I	D	S	$I \vee D$	$I \to S$	$D \to S$	S
1	1	1	**1**	**1**	**1**	**1**
1	1	0	1	0	0	0
1	0	1	**1**	**1**	**1**	**1**
1	0	0	1	0	1	0
0	1	1	**1**	**1**	**1**	**1**
0	1	0	1	1	0	0
0	0	1	0	1	1	1
0	0	0	0	1	1	0

In this example there are three lines, line 1, 3 and 5, in which all premisses are true and, as we can see, in each of these lines also the conclusion is true. So, in each case that all premisses are true, the conclusion is true too. We say that S is a *valid* or *logical consequence* of the premisses $I \vee D$, $I \to S$ and $D \to S$.

Definition 2.5 (Logical or valid consequence).
a) B is a *logical* or *valid consequence* of premisses $A_1, \ldots, A_n$:= in each line of the truth table for $A_1, \ldots, A_n$ and B in which all premisses $A_1, \ldots, A_n$ are 1, also B is 1; in other words, there is no line in the truth table in which all premisses $A_1, \ldots, A_n$ are 1 and at the same time B is 0. **Notation**: $A_1, \ldots, A_n \models B$.
b) Let Γ be a (possibly infinite) set of formulas. B is a *logical* or *valid consequence* of Γ := for each interpretation i, if $i(A) = 1$ for all formulas A in Γ, then also $i(B) = 1$. In other words, each interpretation which is a model of all formulas in Γ is also a model of B. **Notation**: $\Gamma \models B$.

The notion of logical (or valid) consequence is a *semantical* notion: it concerns the truth or falsity, and hence the meaning, of the formulas in question. Notice that in case $n = 0$, i.e., there are no premisses, the definition of $A_1, \ldots, A_n \models B$ reduces to the definition of $\models B$: there is no line in the truth table for B in which B is 0.

Next consider the following argument.

If the weather is nice [N], then John will come [C].
The weather is not nice.

Therefore: John will not come.

We may translate these propositions into the following formulas:

$$\frac{\begin{array}{l} N \to C \\ \neg N \end{array}}{\neg C}$$

Again, for convenience, we have used the symbols N and C instead of the atomic formulas P_1, P_2 in order to help our memory.

Intuitively, this argument is not correct: John may also come when the weather is not nice; for instance, because someone offers John to bring him by car. So, the premisses may be true, while the conclusion is false. We see this clearly in the truth table for the formulas in question:

N	C	$N \to C$	$\neg N$	$\neg C$
1	1	1	0	0
1	0	0	0	1
0	1	**1**	**1**	**0**
0	0	**1**	**1**	**1**

There are two lines in the truth table in which both premisses are 1 (true): line 3 and line 4. In line 4 the conclusion $\neg C$ is 1 too, but in line 3 the conclusion is 0! Line 3 is the case that N is 0 and C is 1, i.e., the weather is not nice, while John does come; in this case both premisses $N \to C$ and $\neg N$ are true, while the conclusion $\neg C$ is false. So, there is a line in the truth table, in which all premisses are true, while the conclusion is false; in other words, $\neg C$ is not a logical consequence of $N \to C$ and $\neg N$. Therefore, not $N \to C, \neg N \models \neg C$ or $N \to C, \neg N \not\models \neg C$.

Notation: Instead of 'not $A_1, \ldots, A_n \models B$' one usually writes: $A_1, \ldots, A_n \not\models B$.

Another intuitive counterexample is the following one; Suppose Berta is a cow and interpret N as 'Berta is a dog' and C as 'Berta has four legs'. Then we have the situation of line 3 in the table: N is 0, C is 1, $N \to C$ is 1, $\neg N$ is 1, but $\neg C$ is 0.

Theorem 2.4.

a) $A \models B$ *if and only if (iff)* $\models A \to B$.

More generally,

b) $A_1, A_2 \models B$ *if and only if (iff)* $A_1 \models A_2 \to B$
if and only if (iff) $\models A_1 \to (A_2 \to B)$
if and only if (iff) $\models A_1 \wedge A_2 \to B$.

Even more generally,

c) $A_1, \ldots, A_n \models B$ *if and only if (iff)* $A_1, \ldots, A_{n-1} \models A_n \to B$
if and only if (iff) $\models (A_1 \wedge \ldots \wedge A_n) \to B$.

Proof. a) $A \models B$ iff there is no line in the truth table in which A is 1 and B is 0. This is equivalent to: there is no line in the truth table in which $A \to B$ is 0. In other words, equivalent to: $\models A \to B$.

b) $A_1, A_2 \models B$ iff there is no line in the truth table in which A_1 and A_2 are both 1 and B is 0. This is equivalent to: there is no line in the truth table in which A_1 is 1 and $A_2 \to B$ is 0, i.e., $A_1 \models A_2 \to B$. This is - in its turn - equivalent to: there is no line in the truth table in which $A_1 \to (A_2 \to B)$ is 0, i.e., $\models A_1 \to (A_2 \to B)$. Or equivalently, there is no line in the truth table in which $(A_1 \wedge A_2) \to B$ is 0, i.e., $\models (A_1 \wedge A_2) \to B$.

c) Similarly. □

It is important to notice that $A \rightarrow B$ is a formula of the logical language, while $\models A \rightarrow B$, or equivalently $A \models B$, is a statement in the meta-language *about* the formulas A and B, namely, that there is no line in the truth table in which A is 1 and B is 0. The symbol $\models$ does not occur in the logical language, but is just an abbreviation from the metalanguage.

Definition 2.6. In the statement $A_1, \ldots, A_n \models B$ we call $A_1, \ldots, A_n$ the *premisses* and B the (putative) *conclusion*. In particular, in $A \models B$ we call A the premiss and B the conclusion. However, in the formula $A \rightarrow B$ we call A the *antecedent* and B the *succedent*.

Theorem 2.5. * *Let Γ be a (possibly infinite) set of formulas. B is a valid consequence of Γ ($\Gamma \models B$) if and only if there are finitely many formulas $A_1, \ldots, A_n$ in Γ such that B is a valid consequence of $A_1, \ldots, A_n$ ($A_1, \ldots, A_n \models B$).*

Proof. The 'if' part is evident. To show the 'only if' part, suppose that $\Gamma \models B$, that is, $\Gamma \cup \{\neg B\}$, i.e., the set consisting of $\neg B$ and of all formulas in Γ, does not have a model. Then, according to the Compactness Theorem 2.3, there is a finite subset $\Gamma' = \{A_1, \ldots, A_n\}$ of formulas in Γ such that $\{A_1, \ldots, A_n\} \cup \{\neg B\}$ does not have a model, which means that $A_1, \ldots, A_n \models B$. □

2.3.1 Decidability

The notion of validity (for the classical propositional calculus) is clearly *decidable*, i.e., there is an algorithm (an effective computational procedure), also called a *decision procedure*, to determine for any formula A in a finite number of steps (depending on the complexity of A) whether it is valid or not. Namely, in order to determine whether A is valid or not, we simply have to compute the truth table of A, entered from its atomic components, and see whether it has 1 in all its lines or not. Computing a truth table of a given formula A and checking whether it has 1 in all its lines can be carried out by a machine and yields an answer 'yes' or 'no' in finitely many steps, the number of steps depending on the complexity of A. Because $A_1, \ldots, A_n \models B$ is equivalent to $\models A_1 \wedge \ldots \wedge A_n \rightarrow B$, also the notion of valid consequence (of a finite number of premisses) is clearly decidable.

One of Leibniz' ideals was to develop a *lingua philosophica* or *characteristica universalis*, an artificial language that in its structure would mirror the structure of thought and that would not be affected with ambiguity and vagueness like ordinary language. His idea was that in such a language the linguistic expressions would be pictures, as it were, of the thoughts they represent, such that signs of complex thoughts are always built up in a unique way out of the signs for their composing parts. Leibniz (1646 - 1716) believed that such a language would greatly facilitate thinking and communication and that it would permit the development of mechanical rules for deciding all questions of consistency or consequence. The language, when it is perfected, should be such that 'men of good will desiring to settle a controversy on any subject whatsoever will take their pens in their hands and say *Calculemus* (let us calculate)'. If we restrict ourselves to the propositional calculus, Leibniz' ideal has been realized: the classical propositional calculus is decidable, more precisely, given premisses $A_1, \ldots, A_n$ and a putative conclusion B, one may decide whether B is a logical consequence of $A_1, \ldots, A_n$ by simply calculating the truth tables of $A_1, \ldots, A_n, B$. However, A. Church and A. Turing proved in 1936 that the predicate calculus is undecidable, i.e., there is no mechanical method to test logical consequence (in the predicate calculus), let alone philosophical truth.

For more information the reader is referred to W. & M. Kneale [16], *The Development of Logic* and to B. Mates [20], *Elementary Logic*, Chapter 12.

Now, if A has been built from n atomic formulas, the truth table of A has 2^n lines. So, a formula built from 10 atomic formulas has a truth table with $2^{10} = 1024$ lines. And if $n = 20$, the truth table of A has $2^{20} = 2^{10} \times 2^{10} = 1024 \times 1024$, so more than a million lines. Hence, the number of steps needed to decide whether a given formula A is valid or not grows fast if A becomes more complex. In fact, if A has been built from 64 atomic formulas, it will take many lifetimes in order to compute whether A is valid or not, even with very futuristic computers, the number of lines being $2^{64} = 2^4 \times (2^{10})^6 \approx 16 \times (10^3)^6 = 16 \times 10^{18}$. In Subsection 2.5.3 we will construct such a formula, built from 64 atomic formulas, to describe a particular travelling salesman problem. Supposing a computer computes $16000 = 16 \times 10^3$ lines per second, in one human lifetime it can compute about 100 (years) $\times$ 365 (days) $\times$ 24 (hours) $\times$ 60 (minutes) $\times$ 60 (seconds) $\times$ 16000 (lines) $\approx 16 \times 10^{13}$ lines. So, in order to compute a truth table of a formula built from 64 atomic formulas, our computer needs about $(16 \times 10^{18}) / (16 \times 10^{13}) = 10^5$ human lifetimes, supposing it can compute 16000 lines per second. This means that our decision procedure to determine whether a given formula A (of the propositional calculus) is valid or not, is a rather theoretical one if the complexity of A is great, more precisely, if A has been built from say 64 atomic components.

One may wonder whether there are more effective or more realistic decision procedures to determine validity, other than making the truth table and checking whether it has 1 in all its lines. No such procedure is known, although for many concrete formulas ad hoc solutions can give a quick answer to the question whether they are valid or not. But no (general) procedure is known, other than making truth tables, to determine the validity of an arbitrary formula.

2.3.2 Sound versus Plausible Arguments; Enthymemes

A concrete argument consists of a number of premisses and a (putative) conclusion. The atomic propositions of the argument are translated into atomic formulas $P_1, P_2, \ldots$ and the composite propositions of the argument are translated into composite formulas which are composed by the logical connectives from the atomic formulas. The result is a logical reasoning pattern:

premisses
|
logical | reasoning
|
conclusion

A reasoning pattern is *valid* if it is impossible that the premisses are true and at the same time the conclusion false. A concrete argument is *correct* if the underlying reasoning pattern is valid, otherwise it is incorrect.

The correctness of a concrete argument is not determined by the content or meaning of the atomic propositions in question, but by the meaning of the propositional connectives (and in predicate logic also by the meaning of the quantifiers) which occur in the argument. That is why one abstracts from the content of the atomic propositions in question by translating them into $P_1, P_2, \ldots$, as pointed out by Frege [8] in his Begriffsschrift (1879).

The atomic formulas may be interpreted as true or false propositions, denoted by 1 and 0 respectively, and the meaning of the logical connectives is specified precisely in the truth tables. Validity of a reasoning pattern means that for every interpretation of the atomic formulas it is impossible that the premisses become true propositions while the conclusion becomes a false proposition.

In his Begriffsschrift [8] of 1879 Gottlob Frege compares the use of the logical language with the use of a microscope. Although the eye is superior to the microscope, for certain distinctions the microscope is more appropriate than the naked eye. Similarly, although natural language is superior to the logical language, for judging the correctness of a certain argument the logical language is more appropriate than natural language. Since the content or meaning of the atomic propositions does not matter for the correctness of the argument, it is more convenient to abstract from this content by replacing the atomic propositions by atomic formulas $P_1, P_2, \ldots$.

It is possible that the study of logic does not augment our native capacity to *discover* correct arguments; but it certainly is of value in *checking* the correctness of given arguments. However, the reader should realize that at this stage we are not yet able to give an adequate logical analysis of, for instance, the following argument.

All men are mortal.
Socrates is a man.
Therefore: Socrates is mortal.

In order to see the correctness of this argument one has to take into account the internal subject-predicate structure of the atomic propositions involved, and this is precisely what is ignored in the propositional calculus and what we shall study in the predicate calculus; see Chapter 4. Using only the means of the propositional calculus, all we can say is that the foregoing argument is of the form $P, Q \models R$, which does not hold, because we may interpret P and Q as true propositions and R as a false one; in other words, P and Q may have the value 1, while R may have the value 0. In order to see the correctness of the argument above, one has to analyse the internal subject-predicate structure of the atomic formulas P, Q and R; but this is beyond the scope of the propositional calculus. In the propositional calculus we can adequately analyse only those arguments the correctness of which depends on the way the composite propositions are composed of the atomic propositions by means of the propositional operations.

Arguments are frequently used to persuade the hearer of the truth of the conclusion on the grounds that (i) the conclusion logically follows from the premisses and in addition (ii) the premisses are true. Let us use $A_1, \ldots, A_n :: B$ to denote
(i) $A_1, \ldots, A_n \models B$, and
(ii) $A_1, \ldots, A_n$ are true; and therefore B is true.

When both (i) and (ii) hold, we call the argument not simply 'valid', but *sound*. And we call an argument *plausible*, when it is valid, but we can only say that $A_1, \ldots, A_n$ are plausible.

It frequently happens that speakers in giving an argument do not explicitly mention *all* their premisses; in some cases they even leave the conclusion tacit. For instance, if someone offers me coffee, I might respond as follows:

If I drink coffee [C], I can't get to sleep early [$\neg S$]. So please don't pour me any.

The argument given is of the form $C \to \neg S :: \neg C$, which is clearly an abbreviation for $C \to \neg S,\ S :: \neg C$.

I might even leave out the conclusion; if I have just been offered a cup of coffee, simply $C \to \neg S$ might be sufficient not to let the hostess pour me any coffee.

Arguments in which one or more premisses or the conclusion is tacit are called *enthymemes*. Premises may not be explicitly stated for practical reasons, but also to mislead the audience.

Exercise 2.19. Translate the propositions in the following argument into formulas of the language of propositional logic and check whether the (putative) conclusion is a logical (or valid) consequence of the premisses:
If the government raises taxes for its citizens, the unemployment grows.
The unemployment does not grow or the income of the state decreases.
Therefore: if the government raises taxes, then the income of the state decreases.

Exercise 2.20. Translate the propositions in the following argument into formulas of the language of propositional logic and check whether the putative conclusion is a logical (or valid) consequence of the premisses:
Europe may form a monetary union only if it is a political union.
Europe is not a political union or all European countries are member of the union.
Therefore: If all European countries are a member of the union, then Europe may form a monetary union.

Exercise 2.21. Verify by making truth tables:
a) $A,\ A \to B \models B$ b) $A \to B,\ \neg B \models \neg A$ c) $A, \neg A \models B$
d) $A \to B \not\models B \to A$ e) $A \to B, \neg A \not\models \neg B$ f) $A \to (B \vee C) \models (A \to B) \vee (A \to C)$
g) $A \vee B,\ \neg A \models B$ h) $\neg(A \wedge B),\ A \models \neg B$

Exercise 2.22. Translate the propositions in the following argument into formulas of the language of propositional logic and check whether the putative conclusion is

a logical (or valid) consequence of the premisses:
John does not win the lottery or he makes a journey [J].
If John does not make a journey, then he does not succeed for logic.
John wins the lottery [W] or he succeeds for logic [S].
Therefore: John makes a journey.

Exercise 2.23. Translate the propositions in the following argument into formulas of the language of propositional logic and check whether the putative conclusion is a logical (or valid) consequence of the premisses:
If Turkey joins the EU [T], then the EU becomes larger [L].
It is not the case that the EU becomes stronger [S] and at the same time not larger.
Therefore: Turkey does not join the EU or the EU becomes stronger.

2.4 Semantics: Meta-logical Considerations

In this section we will prove results *about* the notions of validity and valid consequence of the type: if certain formulas are valid, then also some other formulas are valid.

Suppose we want to determine whether the formula $(P_3 \wedge \neg P_4) \wedge (\neg P_4 \vee P_5 \vee P_6) \rightarrow (P_3 \wedge \neg P_4)$ is valid. Making the truth table of this formula, starting with the atomic formulas P_3, P_4, P_5, P_6 occurring in it, will yield a positive answer. But this table contains $2^4 = 16$ rows and the chance of making a computational mistake is considerable. However, notice that the formula has the form $P_1 \wedge P_2 \rightarrow P_1$ with P_1 replaced by $A_1 = (P_3 \wedge \neg P_4)$ and P_2 replaced by $A_2 = (\neg P_4 \vee P_5 \vee P_6)$. Although the table for $A_1 \wedge A_2 \rightarrow A_1$ may consist of many lines, 16 in our example, there cannot be more than 4 different combinations of 1 and 0 for A_1 and A_2. In our example the second row, in which $A_1 = P_3 \wedge \neg P_4$ has value 1 and $A_2 = \neg P_4 \vee P_5 \vee P_6$ has value 0, will even not occur, because if $\neg P_4$ is 1, then also $A_2 = \neg P_4 \vee P_5 \vee P_6$ is 1.

A_1	A_2	$A_1 \wedge A_2 \rightarrow A_1$
1	1	$(1 \wedge 1) \rightarrow 1 = 1$
1	0	$(1 \wedge 0) \rightarrow 1 = 1$
0	1	$(0 \wedge 1) \rightarrow 0 = 1$
0	0	$(0 \wedge 0) \rightarrow 0 = 1$

All four possible combinations of 1 and 0 for A_1 and A_2 will yield for $A_1 \wedge A_2 \rightarrow A_1$ the value 1. So, from the fact that the formula $P_1 \wedge P_2 \rightarrow P_1$ is valid, we may conclude that also the formula $A_1 \wedge A_2 \rightarrow A_1$ is valid for any formulas A_1 and A_2; in particular, that the formula $(P_3 \wedge \neg P_4) \wedge (\neg P_4 \vee P_5 \vee P_6) \rightarrow (P_3 \wedge \neg P_4)$ is valid. What we have won is that the table for $P_1 \wedge P_2 \rightarrow P_1$ requires only the computation of 4 instead of 16 rows.

The substitution theorem below reduces the amount of work needed to establish the validity of certain formulas.

Theorem 2.6 (Substitution theorem). *Let $E(P_1,P_2)$ be a formula containing only the atomic formulas P_1,P_2, and let $E(A_1,A_2)$ result from $E(P_1,P_2)$ by substituting formulas A_1,A_2 simultaneously for P_1,P_2, respectively.*

$$\textit{If} \models E(P_1,P_2), \textit{ then} \models E(A_1,A_2).$$

More generally: if $\models E(P_1,\dots,P_n)$, then $\models E(A_1,\dots,A_n)$, where the latter formula results from the former one by replacing the atomic formulas $P_1,\dots,P_n$ by the (composite) formulas $A_1,\dots,A_n$.

So, since $\models P_1 \to P_1$, the substitution theorem tells us that

$\models P_2 \wedge \neg P_3 \to P_2 \wedge \neg P_3$ $\quad (A_1 = P_2 \wedge \neg P_3)$

$\models (P_3 \to P_5 \wedge \neg P_7) \to (P_3 \to P_5 \wedge \neg P_7)$ $\quad (A_1 = P_3 \to P_5 \wedge \neg P_7)$

and so on. So, the purpose of the substitution theorem is to reduce the amount of work needed to establish the validity of certain formulas.

Proof. Suppose $E = E(P_1,\dots,P_n)$ contains only the atomic formulas $P_1,\dots,P_n$ and $\models E$, i.e., the truth table of E entered from the atomic formulas $P_1,\dots,P_n$ is 1 in each line.

P_1	...	P_n	...	E
1	...	1	...	1
⋮		⋮		⋮
0	...	0	...	1

Now $E^* = E(A_1,\dots,A_n)$ results from E by substituting the formulas $A_1,\dots,A_n$ for the atomic formulas $P_1,\dots,P_n$ in E. Let us suppose that the formulas $A_1,\dots,A_n$ and hence also E^* are built from the atomic formulas $Q_1,\dots,Q_k$. Then the computation of the truth table of E^* is as follows.

Q_1	...	Q_k	...	A_1	...	A_n	...	E^*
1	...	1	...		...		...	
⋮		⋮		⋮		⋮		⋮
0	...	0	...		...		...	

Since the construction of E^* from $A_1,\dots,A_n$ is the same as the construction of E from $P_1,\dots,P_n$, the truth table of E^* is computed from those of $A_1,\dots,A_n$ in precisely the same manner as the truth table of E is computed from those of $P_1,\dots,P_n$. Hence, because by assumption the computation of the values of E from the values for $P_1,\dots,P_n$ only yield 1's, also the computation of the values of E^* from the values for $A_1,\dots,A_n$ will only yield 1's. I.e., $\models E^*$.

Note that it may happen that some combinations of 0's and 1's for $A_1,\dots,A_n$ do not occur. For instance, if $A_1 = Q_1 \vee \neg Q_1$, then A_1 will have the value 1 in all lines and the value 0 for A_1 will not occur. □

Remark 2.1. : The converse of the substitution theorem, if $\models E^*$, then $\models E$, does not hold. For instance, let $E(P_1) = P_1$ and let $A_1 = P_2 \to P_2$. Then $E^* = E(A_1) = P_2 \to P_2$ is valid, but $E(P_1) = P_1$ is not valid.

In the next theorem the validity of many formulas is shown by means of the substitution theorem. For example, 5b) says that for any choice of formulas A and B, $B \to A \vee B$ is valid. Taking $A = P_1 \wedge \neg P_2$ and $B = P_2 \to P_3$, we find that $(P_2 \to P_3) \to ((P_1 \wedge \neg P_2) \vee (P_2 \to P_3))$ is valid. This method of proving the validity of the latter formula is much more economical than proving the validity directly from its definition by making the truth table of the latter formula entered from the atomic components P_1, P_2 and P_3; this table would consist of eight lines!

Theorem 2.7. *For any choice of formulas A, B, C:*

1	$\models A \to (B \to A)$	*or* $A \models B \to A$ *or* $A, B \models A$
2	$\models (A \to B) \to ((A \to (B \to C)) \to (A \to C))$	*or* $A \to B, A \to (B \to C), A \models C$
3	$\models A \to (B \to A \wedge B)$	*or* $A, B \models A \wedge B$
4a	$\models A \wedge B \to A$	*or* $A \wedge B \models A$
4b	$\models A \wedge B \to B$	*or* $A \wedge B \models B$
5a	$\models A \to A \vee B$	*or* $A \models A \vee B$
5b	$\models B \to A \vee B$	*or* $B \models A \vee B$
6	$\models (A \to C) \to ((B \to C) \to (A \vee B \to C))$	*or* $A \to C, B \to C \models A \vee B \to C$
7	$\models (A \to B) \to ((A \to \neg B) \to \neg A)$	*or* $A \to B, A \to \neg B \models \neg A$
8	$\models \neg\neg A \to A$	*or* $\neg\neg A \models A$
9	$\models (A \to B) \to ((B \to A) \to (A \rightleftarrows B))$	*or* $A \to B, B \to A \models A \rightleftarrows B$
10a	$\models (A \rightleftarrows B) \to (A \to B)$	*or* $A \rightleftarrows B \models A \to B$
10b	$\models (A \rightleftarrows B) \to (B \to A)$	*or* $A \rightleftarrows B \models B \to A$

Proof. The statements in the right column, after the 'or', are according to Theorem 2.4 equivalent to the corresponding statements in the left column, before the 'or'. The statements in the left column follow from the substitution theorem. For instance, to show 1, $\models A \to (B \to A)$, it is easy to verify that $\models P_1 \to (P_2 \to P_1)$, from which it follows by the substitution theorem that for any formulas A, B, $\models A \to (B \to A)$. □

The student is not expected to learn the list in Theorem 2.7 outright now. In the course of time he or she will become familiar with the most frequently used results.

Later in Section 2.9 it will be shown that all valid formulas may be obtained (or deduced) by applications of Modus Ponens to formulas of the ten forms in Theorem 2.7; this is the so-called *completeness theorem* for propositional logic. For that reason formulas of the form 1, ..., 10 in Theorem 2.7 are called logical *axioms for (classical) propositional logic*. Notice that the formulas in 1 and 2 concern $\to$, the formulas in 3 and 4 concern $\wedge$, the formulas in 5 and 6 concern $\vee$, the formulas in 7 and 8 concern $\neg$ and the formulas in 9 and 10 concern $\rightleftarrows$. For instance, the formulas in 1 and 2 would not be valid if the $\to$ were replaced by any other connective. The completeness theorem says essentially that formulas of these ten forms together characterize the meanings of $\rightleftarrows$, $\to$, $\wedge$, $\vee$ and $\neg$: every valid formula may be obtained by applications of Modus Ponens to formulas of these ten forms.

Paradoxes of Material Implication $\models A \to (B \to A)$, or, equivalently, $A \models B \to A$, and $\models \neg A \to (A \to B)$, or, equivalently, $\neg A \models A \to B$, have been called *paradoxes of* *material implication*. This has been illustrated by examples like the following ones: $A \models B \to A$: I like coffee; therefore: if there is oil in my coffee, I like coffee. $\neg A \models A \to B$: I do not break my legs; therefore: if I break my legs, I will go for skying. This sounds very strange indeed. However, Paul Grice [10] has pointed out that in conversation one is supposed to take social rules into account, such as being relevant and maximally informative. And although $B \to A$ is true when A is true, it is simply misleading to say $B \to A$, or equivalently $\neg B \vee A$, when one knows that A is true, because A is clearly more informative than $B \to A$ or, equivalently, $\neg B \vee A$. Similarly, although $A \to B$ is true when $\neg A$ is true, it is misleading to say $A \to B$, or, equivalently, $\neg A \vee B$, when one has the information $\neg A$, because $\neg A$ is clearly more informative than $A \to B$ or, equivalently $\neg A \vee B$.

Also the proof of the next theorem is by showing that one obtains valid formulas if one replaces A, B, C by the atomic formulas P_1, P_2, P_3; next application of the substitution theorem yields the desired result.

Theorem 2.8. *For any formulas A, B, C:*

11	$\models \neg\neg A \rightleftarrows A$	*law of double negation*
12	$\models A \vee \neg A$	*law of excluded middle*
13	$\models \neg(A \wedge \neg A)$	*law of non-contradiction*
14	$\models \neg A \to (A \to B)$ *or* $\neg A, A \models B$	*ex falso sequitur quod libet*
15	$\models (A \to B) \to ((B \to C) \to (A \to C))$	*or* $A \to B,\ B \to C \models A \to C$

From the table for $\rightleftarrows$ follows immediately the next theorem.

Theorem 2.9. *Let A, B be any formulas. $\models A \rightleftarrows B$ if and only if A and B have the same truth table.*

Proof. Suppose $\models A \rightleftarrows B$. Then from the table for $\rightleftarrows$ it follows that it is impossible that in some line of the truth table one of A, B is 1 while the other is 0. Conversely, suppose A and B have the same truth table. Then in every line of the truth table both formulas are 1 or both formulas are 0. In either case $A \rightleftarrows B$ is 1. Since this holds for every line in the truth table, $\models A \rightleftarrows B$. □

Theorem 2.10. *For any formulas A, B, C:*

16	$\models (A \to B) \ \rightleftarrows\ (\neg B \to \neg A)$	*contraposition*
17a	$\models \neg(A \vee B) \ \rightleftarrows\ \neg A \wedge \neg B$	*De Morgan's laws 1847*
17b	$\models \neg(A \wedge B) \ \rightleftarrows\ \neg A \vee \neg B$	
18	$\models \neg(A \to B) \ \rightleftarrows\ A \wedge \neg B$	
19	$\models (A \rightleftarrows B) \ \rightleftarrows\ (A \to B) \wedge (B \to A)$	
20	$\models A \to B \ \rightleftarrows\ \neg(A \wedge \neg B)$	
21	$\models A \to B \ \rightleftarrows\ \neg A \vee B$	
22	$\models A \wedge (B \vee C) \ \rightleftarrows\ (A \wedge B) \vee (A \wedge C)$	*distributive law*
23	$\models A \vee (B \wedge C) \ \rightleftarrows\ (A \vee B) \wedge (A \vee C)$	*distributive law*
24	$\models A \to (B \to C) \ \rightleftarrows\ B \to (A \to C)$	
25	$\models A \to (B \to C) \ \rightleftarrows\ A \wedge B \to C$	

Proof. One easily verifies that $A \to B$ and $\neg B \to \neg A$ have the same truth truth table. Hence, by theorem 2.9 it follows that $\models (A \to B) \rightleftarrows (\neg B \to \neg A)$. Another way of showing this is to verify that $\models (P_1 \to P_2) \rightleftarrows (\neg P_2 \to \neg P_1)$, simply by computing the truth table. Next the substitution theorem 2.6 yields the desired result.
The other items are shown similarly. □

A reasoning rule like Modus Ponens or Modus Tollens should, of course, be *sound*, i.e., if its premisses are true (1), then its conclusion must be true (1) too. In other words, these rules should preserve truth. One easily verifies that Modus Ponens and Modus Tollens are sound.

Theorem 2.11. *(a) For every line in the truth table: if A is 1 and $A \to B$ is 1 in that line, then B is also 1 in that line. In other words: $A,\ A \to B \models B$.*
We say that the rule of Modus Ponens (MP) is sound. Consequently:
(b) For all formulas A and B, if $\models A$ and $\models A \to B$, then $\models B$. In other words: for all formulas A and B, if $\models A \to B$, then if (in addition) $\models A$, then $\models B$.
(c) However, not for all formulas A and B, if (if $\models A$, then $\models B$), then $\models A \to B$.

Proof. (a) follows immediately from the truth table for $\to$.
From (a) follows: if A is 1 in all lines and $A \to B$ is 1 in all lines of the truth table, then B is 1 in all lines of the truth table. In other words, if $\models A$ and $\models A \to B$, then $\models B$. This proves (b).
(c) 'if $\models A$, then $\models B$' means: if A is 1 in all lines of the truth table, then B is 1 in all lines of the truth table (*). $\models A \to B$ means: in every line in which A is 1, B must be 1 too. Notice that this does *not* follow from (*). For suppose that A is 1 in *some* line of the truth table, we do not know whether A is 1 in *all* lines of its truth table. In fact, there are formulas A and B such that 'if $\models A$, then $\models B$' holds, while $\models A \to B$ does not hold. For example, take $A = P_1$ (it is cold) and $B = P_2$ (it is snowing). Since $\not\models P_1$ (not always it is cold) and $\not\models P_2$ (not always it is snowing), 'if $\models P_1$, then $\models P_2$' holds, while $\models P_1 \to P_2$ (always if it is cold, then it is snowing) does not hold. □

Theorem 2.12. *(a) For all formulas A, if $\models \neg A$, then not $\models A$.*
However, the converse does not hold:
(b) Not for all formulas A, if not $\models A$, then $\models \neg A$.

Proof. (a) Suppose $\models \neg A$, i.e., $\neg A$ is 1 in all lines of its truth table. Equivalently: A is 0 in all lines of its truth table. So, for sure, it is not the case that A is 1 in all lines of its truth table, i.e., not $\models A$.
(b) 'Not $\models A$' means that not in all lines of its truth table A is 1, in other words, A is 0 in *some* line of its truth table. This does not mean that $\models \neg A$, or equivalently, that A is 0 in *all* lines of its truth table. In fact, there are formulas A such that not $\models A$, while $\models \neg A$ does not hold. For instance, take $A = P_1$ (it is raining). Then not $\models P_1$ (not always it is raining), while $\models \neg P_1$ (always it is not raining; it never rains) does not hold. □

Warning One might be inclined to write: for all formulas A, if not $\models A$, then *not* $\models \neg A$. However, this is false. For instance, taking $A = P_1 \wedge \neg P_1$ we have
not $\models P_1 \wedge \neg P_1$, but also $\models \neg(P_1 \wedge \neg P_1)$. The expression

$$\text{if not} \models A, \text{ then} \models \neg A \qquad (*)$$

does hold for some formulas, for instance, for $A = P_1 \wedge \neg P_1$, but it does not hold for other formulas, for instance, not for $A = P_1$.

A formula that refutes (*) is called a *counterexample* to the statement (*). So, P_1 is a counterexample to (*).

Theorem 2.13. *(a) For all formulas A and B, if* $\models A$ *or* $\models B$*, then* $\models A \vee B$*.*
However, the converse does not hold:
(b) Not for all formulas A and B, if $\models A \vee B$*, then* $\models A$ *or* $\models B$*.*

Proof. (a) Suppose $\models A$ or $\models B$. Consider the case that $\models A$, i.e., A is 1 in all lines of its truth table. Then clearly, also $A \vee B$ is 1 in all lines of its truth table, i.e., $\models A \vee B$. The case that $\models B$ is treated similarly.
(b) $\models A \vee B$ means: $A \vee B$ is 1 in all lines of its truth table, i.e., in each line of the truth table A is 1 or B is 1. So, there might be lines in which A is 1 and B is 0, while there might be other lines in which A is 0 and B is 1. So, this does not mean that A is 1 in all lines, i.e., $\models A$, nor that B is 1 in all lines, i.e., $\models B$. In fact, there are formulas A and B, such that $\models A \vee B$, while neither $\models A$ nor $\models B$. For instance, take $A = P_1$ (it is raining) and $B = \neg P_1$. Then $\models P_1 \vee \neg P_1$ (always it is raining or not raining), while neither $\models P_1$ (always it is raining), nor $\models \neg P_1$ (always it is not raining; it never rains). □

Warning One might be inclined to write: for all formulas A and B, if $\models A \vee B$, then not $\models A$ and not $\models B$. However, this is false. For instance, take $A = P_1 \to P_1$ and B arbitrary, then $\models (P_1 \to P_1) \vee B$, but also $\models P_1 \to P_1$ holds. The expression

$$\text{if} \models A \vee B, \text{ then} \models A \text{ or} \models B \qquad (*)$$

does hold for some formulas, for instance, for $A = P_1 \to P_1$ and B arbitrary, but it does not hold for other formulas, for instance, not for $A = P_1$ and $B = \neg P_1$. So, $A = P_1$ and $B = \neg P_1$ is a *counterexample* to the statement (*).

Notice that, for instance, $A = P$ and $B = Q$ with P, Q atomic, is not a counterexample against (*), because such a counterexample should consist of formulas A and B such that '$\models A \vee B$' does hold, while '$\models A$ or $\models B$' does not hold; and $\models P \vee Q$ is not the case.

Theorem 2.14. *For all formulas A and B,* $\models A \wedge B$ *if and only if* $\models A$ *and* $\models B$*.*

Proof. $\models A \wedge B$ means: in all lines of its truth table, $A \wedge B$ is 1, i.e., in all lines, A is 1 and B is 1. This is equivalent to: in all lines A is 1 and in all lines B is 1, i.e., $\models A$ and $\models B$. □

In order to be able to formulate the replacement theorem, we first have to define the notion of *subformula*.

Definition 2.7 (Subformula). 1. If A is a formula, then A is a subformula of A.
2. If A and B are formulas, the subformulas of A and the subformulas of B are subformulas of $A \rightleftarrows B$, $A \to B$, $A \wedge B$, and $A \vee B$.
3. If A is a formula, then the subformulas of A are subformulas of $\neg A$.

Example 2.6. The subformulas of $\neg P \vee Q \to (P \to \neg P \vee Q)$ are: $\neg P \vee Q \to (P \to \neg P \vee Q)$, $\neg P \vee Q$, $P \to \neg P \vee Q$, $\neg P$, Q and P. Notice that $P \vee Q$ is not a subformula of $\neg P \vee Q \to (P \to \neg P \vee Q)$.

Theorem 2.15 (Replacement theorem). *Let C_A be a formula containing A as a subformula, and let C_B come from C_A by replacing the subformula A by formula B. If A and B have the same truth table, then C_A and C_B have the same truth table too.*

Proof. Assume A and B have the same table. If, in the computation of a given line of the table for C_A, we replace the computation of the specified part A by a computation of B instead, the outcome will be unchanged. Thus, C_B has the same table as C_A. □

Corollary 2.1 (Replacement rule). *If $\models C_A$ and A and B have the same table, then $\models C_B$.*

Warning: do not confuse object- and meta-language The reader should realize that the symbol '$\models$' does not occur in the alphabet of the propositional calculus and that consequently any expression containing $\models$ is not a formula. '$\models A$' is a statement *about* formula A, saying that A is valid, i.e., A is 1 in all lines of its truth table (always true). 'A' stands for a formula in the *object-language*, i.e., the language of propositional logic, but '$\models A$' is an expression in the *meta-language* about formula A, saying that A is always true.

$\models A \rightleftarrows B$ means $\models (A \rightleftarrows B)$; it cannot mean $(\models A) \rightleftarrows B$, because '$\models A$' belongs to the meta-language, while '$\rightleftarrows$' and 'B' belong to the object language. So, '$\models$' stands outside every formula.

Because '$\models \neg A$' is an expression of the meta-language and '$\to$' is a symbol of the object language, we are not allowed to write 'if $\models \neg A$, then not $\models A$' in Theorem 2.12 as '$\models \neg A \to \neg \models A$'; '$\to$' should connect formulas and '$\models \neg A$' and 'not $\models A$' are not formulas.

We can compare '$\models A$' with for instance "'Jean est malade' is a short sentence". This is not a sentence in French (the object language), but a statement in English (the meta-language) about a sentence ('Jean est malade', 'A') of the object language.

Below we have listed a number of expressions on the left hand side and the language they belong to on the right hand side.

$P \wedge \neg P$:	Formula of the object-language.
$\models P \wedge \neg P$:	Statement in the meta-language about the formula $P \wedge \neg P$.
'$\models P \wedge \neg P$' is false:	Statement in the meta-meta-language about $\models P \wedge \neg P$.

Because our meta-language is a natural language (English), the meta-meta-language coincides with the meta-language itself.

Exercise 2.24. Show that for all formulas A and B,
1) if $\models A \rightleftarrows (A \to B)$, then $\models A$ and $\models B$;
2) if $A \models \neg A$, then $\models \neg A$.
3) if $A \to B \models A$, then $\models A$.

Exercise 2.25. Prove or refute: for all formulas A and B,

a) if not $\models A \to B$, then $\models A$ and $\models \neg B$.
b) if $\models \neg(A \to B)$, then $\models A$ and $\models \neg B$.
c) if not $\models A \wedge B$, then $\models \neg A$ or $\models \neg B$.
d) if $\models \neg(A \wedge B)$, then $\models \neg A$ or $\models \neg B$.
e) if not $\models A \vee B$, then $\models \neg A$ and $\models \neg B$.
f) if $\models \neg(A \vee B)$, then $\models \neg A$ and $\models \neg B$.

Exercise 2.26. Establish the following.
(a1) $A_1, A_2, A_3 \models A_1,\ A_1, A_2, A_3 \models A_2,\ A_1, A_2, A_3 \models A_3$.
(a2) More generally: $A_1, \ldots, A_i, \ldots, A_n \models A_i$ for $i = 1, \ldots, n$.
(b1) If $A_1, A_2, A_3 \models B_1$ and $A_1, A_2, A_3 \models B_2$ and $B_1, B_2 \models C$, then $A_1, A_2, A_3 \models C$.
(b2) More generally, for any $n, k \geq 0$: if $A_1, \ldots, A_n \models B_1$ and ... and $A_1, \ldots, A_n \models B_k$ and $B_1, \ldots, B_k \models C$, then $A_1, \ldots, A_n \models C$.

Exercise 2.27. Show directly from the definition of valid consequence:
1) if $A \models B$ and $A \models \neg B$, then $\models \neg A$. (*Reductio ad absurdum*)
2) if $A \models C$ and $B \models C$, then $A \vee B \models C$. (*Proof by cases*)

Exercise 2.28. Which of the following statements are right and which are wrong, and why is that the case? For all formulas A, B, C,
(a) $A \to B \vee C \models (A \to B) \vee (A \to C)$.
(b) if $\models (A \to B) \vee (A \to C)$, then $\models A \to B$ or $\models A \to C$.
(c) if $A \models B$, then $B \to C \models A \to C$.

Exercise 2.29. Prove: if $T \wedge A \wedge B \models P$, then $\neg P \models \neg T \vee \neg A \vee \neg B$.
Interpreting T as a Theory, A as Auxiliary hypotheses, B as Background hypotheses and P as Prediction, this is actually the *Duhem-Quine thesis*. In 1906 Pierre Duhem argued that the falsification of a theory is necessarily ambiguous and therefore that there are no crucial experiments; one can never be sure that it is a given theory rather than auxiliary or background hypotheses which experiment has falsified. [See S.C. Harding, [11], *Can theories be refuted*? p. IX.]

Exercise 2.30. Prove or refute: for all formulas A, B and C,
a) if $A \models B$, then $\neg B \models \neg A$.
b) if $A \models B$ and $A,\ B \models C$, then $A \models C$.
c) if $A \vee B \models A \wedge B$, then A and B have the same truth table.

2.5 About Truthfunctional Connectives

One may wonder if the object-language of propositional logic may be enriched by adding some new truthfunctional connectives, for instance, the connective $\uparrow$, called the *Sheffer stroke*, to be read as 'neither ..., nor ...' and defined by the following truth table.

A	B	$A \uparrow B$
1	1	0
1	0	0
0	1	0
0	**0**	**1**

In this case we see immediately that $\uparrow$ may be defined in terms of $\neg$ and $\wedge$: $A \uparrow B$ has the same truth table as $\neg A \wedge \neg B$. But maybe there are other *binary* (i.e., with two arguments A and B) *truthfunctional connectives* which cannot be defined in terms of the ones we already have: $\rightleftarrows$, $\rightarrow$, $\wedge$, $\vee$ and $\neg$.

Now it is easy to see that there are $2^4 = 16$ possible binary truthfunctional connectives, each of them corresponding with a table of length 4:

A	*B*			...		
1	1	1	1	...	0	0
1	0	1	1	...	0	0
0	1	1	1	...	0	0
0	0	1	0	...	1	0

It is not difficult to see that each of these 16 truthfunctional connectives may be expressed in terms of $\wedge$, $\vee$ and $\neg$. Consider, for instance, the three truthfunctional connectives corresponding with the following truth tables:

A	*B*	
1	1	0
1	**0**	**1**
0	1	0
0	0	0

A	*B*	
1	1	0
1	0	0
0	**1**	**1**
0	0	0

A	*B*	
1	1	0
1	**0**	**1**
0	**1**	**1**
0	0	0

The left truth table is precisely the table of $A \wedge \neg B$, the truth table in the middle is precisely the table of $\neg A \wedge B$, and the right truth table is precisely the truth table of $(A \wedge \neg B) \vee (\neg A \wedge B)$. So, the following Theorem is evident:

Theorem 2.16. *Each binary (i.e., having two arguments A and B) truthfunctional connective may be expressed in terms of* $\wedge$, $\vee$ *and* $\neg$.

We say that the set $\{\wedge, \vee, \neg\}$ is a *complete set of truthfunctional connectives*: each binary truthfunctional connective may be expressed in terms of these three connectives. We have already seen earlier that $\rightarrow$ and $\rightleftarrows$ can be expressed in terms of $\wedge$, $\vee$ and $\neg$: $A \rightarrow B$ has the same truth table as $\neg A \vee B$, and also as $\neg(A \wedge \neg B)$; and $A \rightleftarrows B$ has the same truth table as $(A \rightarrow B) \wedge (B \rightarrow A)$.

Theorem 2.16 can easily be generalized to truth tables entered from more than two formulas. Consider, for instance, the truth table below entered from three atomic formulas P, Q and R:

P	*Q*	*R*	
1	1	1	0
1	**1**	**0**	**1**
1	0	1	0
1	0	0	0
0	**1**	**1**	**1**
0	1	0	0
0	0	1	0
0	0	0	0

The formula corresponding with this table is clearly: $(P \wedge Q \wedge \neg R) \vee (\neg P \wedge Q \wedge R)$. More generally, we see that for every formula A there is a formula A' which is a disjunction of conjunctions of *literals*, i.e., atomic formulas or negations of atomic formulas, such that A and A' have the same truth table. We shall say that A' is in *disjunctive normal form*. By applying the de Morgan's laws (Theorem 2.10), we may conclude that for every formula A there is also a formula A'' in *conjunctive normal form*, i.e., which is a conjunction of disjunctions of literals, and which has the same truth table as A. See Theorem 2.18.

Next we shall show that each truthfunctional connective may be expressed in terms of only one connective: the Sheffer stroke $\uparrow$.

Theorem 2.17. *Every binary truthfunctional connective may be expressed in terms of the Sheffer stroke* $\uparrow$.

Proof. In order to prove this, by Theorem 2.16 it suffices to prove that $\wedge$, $\vee$ and $\neg$ may be expressed in terms of the Sheffer stroke $\uparrow$.

a) $\neg A$ has the same truth table as $\neg A \wedge \neg A$, and hence as $A \uparrow A$ (neither A, nor A).
b) $A \wedge B$ has the same truth table as $\neg(\neg A) \wedge \neg(\neg B)$, hence as $\neg A \uparrow \neg B$ (neither $\neg A$, nor $\neg B$) and therefore as $(A \uparrow A) \uparrow (B \uparrow B)$.
c) $A \vee B$ has the same truth table as $\neg(\neg A \wedge \neg B)$, hence as $\neg(A \uparrow B)$ and therefore as $(A \uparrow B) \uparrow (A \uparrow B)$. □

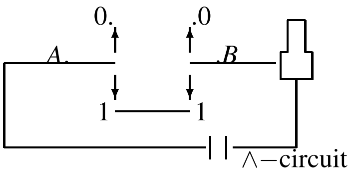
2.5.1 Applications in Electrical Engineering and in Jurisdiction

There are many situations in which there are two opposites analogous to the case of truth and falsity of propositions. For example, in electrical engineering: on (lit, 1) and off (unlit, 0); and in jurisdiction: innocent and guilty. In all such situations one can work with truth tables in a similar way as in propositional logic.

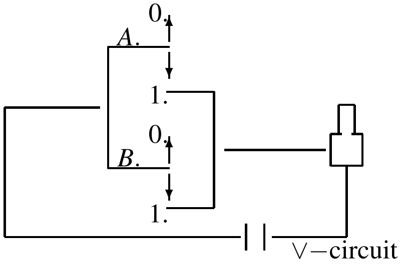
Suppose we have two switches A and B, both with a 0- and a 1- position, a bulb and a battery and that we want the bulb to burn (1, lit) precisely if both switches are in the 1-position. So, the corresponding table is the one for $A \wedge B$:

switch A	switch B	bulb
1	1	1
1	0	0
0	1	0
0	0	0

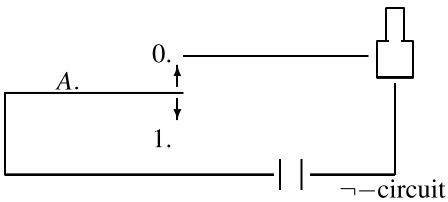
The following electric circuit will satisfy our wishes.



If we want the bulb to burn if at least one of the two switches A and B is in the 1-position, then we find a table corresponding with the one for $A \vee B$ and the corresponding electric circuit is as follows.



And if we want the bulb to burn if switch A is in the 0-position, then we find a table corresponding with the one for $\neg A$ and the corresponding electric circuit is the following one.

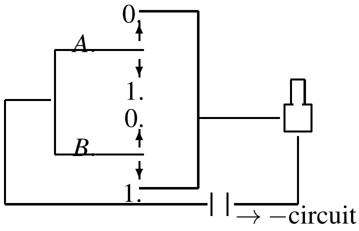


Theorem 2.16 formulated in terms of electric circuits now tells us that each electric circuit can be built from the electric circuits for $\wedge$, $\vee$ and $\neg$, and the proof of Theorem 2.16 provides us with a uniform method to build any circuit we want from the circuits for $\wedge$, $\vee$ and $\neg$. We shall consider some examples below. However, the circuits resulting from our uniform method in the proof of Theorem 2.16 will not always be the simplest ones and for economic reasons one may in practice use circuits other than the ones found by this uniform method.

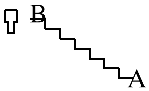
Example 2.7. Suppose we want our bulb to burn in all cases except one: if switch A is in position 1 and switch B is in position 0. So the corresponding table is the following one.

switch A	switch B	bulb
1	1	1
1	0	0
0	1	1
0	0	1

We see that this table corresponds with the one for $A \to B$. The proof of Theorem 2.16 tells us that the circuit corresponding with $(A \wedge B) \vee (\neg A \wedge B) \vee (\neg A \wedge \neg B)$ will satisfy our wishes. However, a much simpler, and hence less expensive circuit, doing the same job, can be found if we realize that $A \to B$ has the same truth table as $(\neg A) \vee B$. So in order to achieve our purpose, we can take the $\vee$-circuit described above with instead of switch A the circuit for $\neg A$.



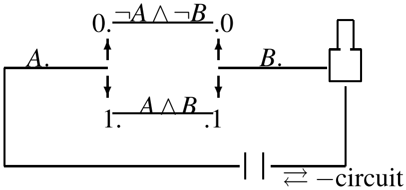
Example 2.8. Suppose we want to build a two-way switch: a switch A at the foot of the stairs and a switch B at the top of the stairs such that we can turn the light on and off both at the foot and at the top of the stairs by changing the nearest switch over into another position.



We can achieve our purpose by making the electric circuit such that the light is on when both switches are in the same position and off when both are in a different position. The corresponding table is the following one.

switch A	switch B	light
1	1	1
1	0	0
0	1	0
0	0	1

This table corresponds with the one for $A \rightleftarrows B$. Applying the proof of Theorem 2.16, we shall find that the circuit corresponding with $(A \wedge B) \vee (\neg A \wedge \neg B)$ will satisfy our requirements. So we can take the $C \vee D$-circuit described above with the circuit for $A \wedge B$ instead of switch C and the circuit for $\neg A \wedge \neg B$ instead of switch D. And this latter circuit is obtained by replacing in the $E \wedge F$-circuit described above switch E by the circuit for $\neg A$ and switch F by the circuit for $\neg B$.



For an application of truth tables in jurisdiction we refer the reader to Exercise 2.31.

2.5.2 *Normal Form***; Logic Programming**

Definition 2.8 (Normal form). A *literal* is by definition an atomic formula or the negation of an atomic formula.

A formula B is in *disjunctive normal form* if it is a disjunction $B_1 \vee \ldots \vee B_k$ of formulas, where each B_i ($1 \leq i \leq k$) is a conjunction $L_1 \wedge \ldots \wedge L_n$ of literals.

A formula B is in *conjunctive normal form* if it is a conjunction $B_1 \wedge \ldots \wedge B_k$ of formulas, where each B_i ($1 \leq i \leq k$) is a disjunction $L_1 \vee \ldots \vee L_n$ of literals.

Example 2.9. So, P_2 and $\neg P_1$ are examples of literals. $(\neg P_1 \wedge P_2) \vee \neg P_3$ is a formula in disjunctive normal form, and $(\neg P_1 \vee \neg P_3) \wedge (P_2 \vee \neg P_3)$ is a formula in conjunctive normal form.

Theorem 2.18 (Normal form theorem). *For each formula A (of classical propositional logic) there are formulas A' and A'' in disjunctive or conjunctive normal form respectively, which have the same truth table as A. In other words, each formula A of classical propositional logic may be written in disjunctive, respectively, conjunctive, normal form.*

Proof. We will use the induction principle (Theorem 2.2) to show that every formula A has the property Φ: there are formulas A' and A'' in disjunctive or conjunctive normal form respectively, which have the same truth table as A. Since all truth-functional connectives can be expressed in terms of $\neg$, $\wedge$ and $\vee$, we may assume that all formulas are built from atomic formulas by means of these three connectives.

1. If A is an atomic formula P, then $A = P$ itself is both in disjunctive and in conjunctive normal form.
2. Suppose $A = \neg B$ and (induction hypothesis) that there are formulas B' and B'' which are in disjunctive or conjunctive normal form respectively, and which are equivalent to B. Then $A = \neg B$ has the same truth table as $\neg B'$, which by the De Morgan's laws, Theorem 2.10, 17, can be rewritten as a conjunction of disjunctions of literals. And $A = \neg B$ has the same truth table as $\neg B''$, which by the De Morgan's laws, Theorem 2.10, 17, can be rewritten as a disjunction of conjunctions of literals.
3. Suppose $A = B \wedge C$ and (induction hypothesis) that there are formulas B', C' and formulas B'', C'' which are in disjunctive or conjunctive normal form respectively and which are equivalent to B, respectively C. Then $A = B \wedge C$ has the same truth table as $B'' \wedge C''$, which is again a conjunction of disjunctions of literals. And $A = B \wedge C$ has the same truth table as $B' \wedge C'$, which by the distributive laws, Theorem 2.10, 22 and 23, can be rewritten in disjunctive normal form.
4. Suppose $A = B \vee C$ and (induction hypothesis) that there are formulas B', C' and formulas B'', C'' which are in disjunctive or conjunctive normal form respectively and which are equivalent to B, respectively C. Then $A = B \vee C$ has the same truth table as $B' \vee C'$, which is again a disjunction of conjunctions of literals. And $A = B \vee C$ has the same truth table as $B'' \vee C''$, which by the distributive laws, Theorem 2.10, 22 and 23, can be rewritten as a conjunction of disjunctions of literals.

Example 2.10. $A = P \rightarrow \neg(\neg Q \vee P)$ has the same truth table as, subsequently, $\neg P \vee \neg(\neg Q \vee P)$, $\neg P \vee (\neg\neg Q \wedge \neg P)$, $\neg P \vee (Q \wedge \neg P)$, which is in disjunctive normal form, and $(\neg P \vee Q) \wedge (\neg P \vee \neg P)$, which is in conjunctive normal form.

Knowledge Representation and Logic Programming The language of logic may be used to represent knowledge. For instance, suppose a person has the following knowledge at his disposal:
(1) John buys the book if it is about logic and interesting.
(2) The book is about logic.
(3) The book is interesting if it is about logic.
Using P to represent 'John buys the book',
Q to represent 'the book is about logic', and
R to represent 'the book is interesting',
the person's knowledge can be represented by the following logical formulas:
(1a) $Q \wedge R \rightarrow P$,
(2a) Q,
(3a) $Q \rightarrow R$.

In the programming language Prolog (Programming in Logic), which will be treated in Chapter 9, these formulas are rendered as follows:
(1b) P :- Q, R. (to be read as: P if Q and R)
(2b) Q.
(3b) R :- Q. (to be read as: R if Q)

(1b) and (3b) are called *rules* and (2b) is called a *fact*. Using logical reasoning 'new' knowledge can be deduced from the knowledge already available. For instance, from (2a) and (3a) follows R (4a), and from (2a), (4a) and (1a) follows P, i.e., 'John buys the book'.

(1b), (2b) and (3b) together can be considered to form a *knowledge base* from which new knowledge can be obtained by logical reasoning or deduction.

The programming language Prolog, to be treated in Chapter 9, has a built in logical inference mechanism. When provided with the database consisting of (1b), (2b) and (3b), Prolog will answer the question '?- P.' with 'yes', corresponding to the fact that P is a logical consequence of (1b), (2b) and (3b).

The following definition introduces some terminology which is used in logic programming and which is needed in Chapter 9.

Definition 2.9 (Literal). a) A *positive literal* is an atomic formula. A *negative literal* is the negation of an atomic formula.
b) A *clause* is a formula of the form $L_1 \vee \ldots \vee L_m$, where each L_i is a literal.

Because clauses are so common in logic programming, it will be convenient to adopt a special clausal notation. In logic programming the clause $\neg P_1 \vee \ldots \vee \neg P_k \vee Q_1 \vee \ldots \vee Q_n$, where $P_1, \ldots, P_k, Q_1, \ldots, Q_n$ are atomic, is denoted by

$$Q_1, \ldots, Q_n \text{ :- } P_1, \ldots, P_k \ (k \geq 0).$$

which stands for $P_1 \wedge \ldots \wedge P_k \rightarrow Q_1 \vee \ldots \vee Q_n$, which has the same truth table as $\neg P_1 \vee \ldots \vee \neg P_k \vee Q_1 \vee \ldots \vee Q_n$.

Theorem 2.18 says that each formula of (classical) propositional logic may be written as a finite conjunction of clauses.

For reasons of efficiency, to be explained in Chapter 9, in Prolog only *Horn clauses* are used, i.e., clauses which contain at most one positive literal, in other words, which are of the form Q :- $P_1, \ldots, P_k$ or of the form :- $P_1, \ldots, P_k$.

(1b), (2b) and (3b) above are examples of Horn-clauses. $Q_1,\ Q_2$:- $P_1,\ P_2,\ P_3$. or equivalently $P_1 \wedge P_2 \wedge P_3 \rightarrow Q_1 \vee Q_2$, is not a Horn clause.

Definition 2.10 (Horn clause).

a) A *definite program clause* is a clause of the form

$$Q \text{ :- } P_1, \ldots, P_k \ (k \geq 0, P_1, \ldots, P_k, Q \text{ atomic})$$

which contains precisely one atomic formula (viz. Q) in its consequent. Q is called the *head* and $P_1, \ldots, P_k$ is called the *body* of the program clause.

b) A *unit clause*, also called a *fact*, is a clause of the form

$$Q \text{ :-}$$

that is, a definite program clause with an empty body.

c) A *definite program* is a finite set of definite program clauses.

d) A *definite goal* is a clause of the form

$$\text{:- } P_1, \ldots, P_k$$

that is, a clause which has an empty consequent. Each P_i $(i = 1, \ldots, k)$ is called a *subgoal* of the goal.

e) A *Horn clause* is a clause which is either a definite program clause or a definite goal. So, a Horn clause is a clause with at most one positive literal.

Example 2.11. The following is an example of a definite program:

P :- Q, R.
Q :-.
R :- Q.

This program corresponds with the formula $(P \vee \neg Q \vee \neg R) \wedge Q \wedge (R \vee \neg Q)$, which is in conjunctive normal form, and where each conjunct contains precisely one positive literal (and hence is a Horn clause). Note that this formula has the same truth table as $(Q \wedge R \rightarrow P) \wedge Q \wedge (Q \rightarrow R)$.

Given this program, in logic programming the goal ':- P.' will be answered with 'yes', corresponding with the fact that P logically follows from $(P \vee \neg Q \vee \neg R) \wedge Q \wedge (R \vee \neg Q)$.

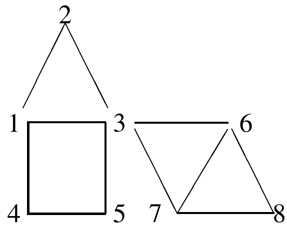
The goal ':- S' will be answered with 'no', corresponding with the fact that S does not logically follow from the given program.

Logic programming in general and Prolog in particular will be treated in Chapter 9. However, this treatment also presupposes familiarity with classical predicate logic, which will be treated in Chapter 4.

2.5.3 *Travelling Salesman Problem (TSP)**; *NP-completeness**

The Traveling Salesman Problem is the problem of computing the shortest itinerary, when a number, n, of cities with given distances has to be visited, each city to be visited only once. From a theoretical point of view there is no problem at all: if there are n cities to be visited, there are $(n-1)!$ itineraries; compute the total distance of each of them and take the shortest. However, from a practical point of view there are problems: if 10 cities are to be visited, there are 9! = 362,880 itineraries; and if a sales-representative has to visit 30 cities, there are 29! itineraries and 29! is larger than 10^{29}. Supposing that a computer could calculate the distances of 1000 = 10^3 itineraries per second, in one human lifetime it could compute about 100 (years) $\times$ 365 (days) $\times$ 24 (hours) $\times$ 60 (minutes) $\times$ 60 (seconds) $\times$ 10^3 (itineraries) $\approx 10^{13}$ itineraries. So, in order to compute the distances of 29! itineraries, our computer would need more than 10^{29} / $10^{13} = 10^{16}$ human lifetimes! Thus, like the validity problem for formulas of propositional logic, also the Travelling Salesman Problem is solvable in theory, but no realistic solution is known.

We will see below how the following *Traveling Salesman Problem* can be reduced to a *satisfiability problem* in the propositional calculus. In the map, the vertices are towns and the lines are roads, each 10 miles long. This example is from A. Keith Austin [1].



PROBLEM: Can the salesman start at 1 and visit all the towns in a journey of only 70 miles?

Theorem 2.19. *There is a formula E of the propositional calculus such that there is a journey of only 70 miles starting at 1 if and only if E is satisfiable.*

CONSTRUCTION of E: To express the problem in propositional logic, we introduce the atomic formulas P_t^m, for $m = 0, 1, \ldots, 7$, $t = 1, 2, \ldots, 8$, the intended meaning of P_t^m being: after $10 \times m$ miles the salesman is at town t. Given any journey of 70 miles, each P_t^m is either true or false. We now express the conditions of the problem as logical formulas.

i) If the salesman is at 5 after 30 miles, then he is at 3 or 4 after 40 miles, i.e., if P_5^3 is true, then either P_3^4 or P_4^4 is true. Let $J_5^3 := P_5^3 \rightarrow P_3^4 \vee P_4^4$ be the formula in our propositional language expressing this. Similarly we have $P_5^m \rightarrow P_3^{m+1} \vee P_4^{m+1}$, and $P_3^m \rightarrow (P_1^{m+1} \vee P_2^{m+1} \vee P_5^{m+1} \vee P_6^{m+1} \vee P_7^{m+1})$ for $m = 0, 1, \ldots, 6$, and so on for each town. Denote each of these by the corresponding J_y^x. All these have to be true and so we write

$$J := J_1^0 \wedge J_2^0 \wedge \ldots \wedge J_8^0 \wedge J_1^1 \wedge J_2^1 \wedge \ldots \wedge J_8^1 \wedge \ldots \wedge J_8^6.$$

ii) Another condition is that each town has to be visited. That town 1 has to be visited can be expressed as $P_1^0 \vee P_1^1 \vee P_1^2 \vee \ldots \vee P_1^7$ and similarly for the other towns. Let

$$V := (P_1^0 \vee \ldots \vee P_1^7) \wedge (P_2^0 \vee \ldots \vee P_2^7) \wedge \ldots \wedge (P_8^0 \vee \ldots \vee P_8^7).$$

iii) Also the salesman is only at one town at any one time, so we have, e.g., $P_3^5 \rightarrow \neg P_1^5$. Let $N_3^5 := P_3^5 \rightarrow \neg P_1^5 \wedge \neg P_2^5 \wedge \neg P_4^5 \wedge \ldots \wedge \neg P_8^5$. And let

$$N := N_1^0 \wedge N_2^0 \wedge \ldots \wedge N_8^7.$$

iv) Finally, he has to start at 1, so we require P_1^0 to be true.

Now let $E := J \wedge V \wedge N \wedge P_1^0$. Then E has the required property: there is a journey of only 70 miles starting at 1 if and only if E is satisfiable.

Theorem 2.19 reduces the Traveling Salesman Problem for eight cities to a satisfiability problem in the propositional calculus. However, the formula E constructed in the proof of Theorem 2.19 is built from $8^2 = 64$ atomic formulas P_t^m. So, in order to check whether E is satisfiable, we have to compute a truth table entered from 2^{64} lines. We have already seen in Subsection 2.3.1 that making truth tables with so many entries does not yield a practical or realistic decision method to decide whether arbitrary formulas are satisfiable or not. Since the original problem can be solved by computing the distances of (8 - 1)! itineraries, the reduction of the Traveling Salesman Problem to the satisfiability problem for propositional logic has not helped us to find a practical or realistic solution for the former. We have to wait for a realistic solution of the satisfiability problem or for a proof that no such solution exists.

Of course, in order to see whether a given formula E is satisfiable, i.e., has at least one 1 in its truth table, one might *non-deterministically* choose a line in the truth table and compute whether E is 1 in that line. The computation of one line in the truth table can be done in a realistic way: the time required to do so is a *polynomial* of the complexity of the formula in question. If it turns out that E is 1 in the chosen line, one knows that E is satisfiable, but when it turns out that E is 0 in the chosen line, one does not know whether E is satisfiable or not. And we have seen in Subsection 2.3.1 that it is not realistic to compute all lines in the truth table of E if E has been built from many, say 64, atomic formulas. For that reason, the satisfiability problem for propositional calculus is said to belong to the class *NP* of all problems which may be decided Non-deterministically in Polynomial time.

In 1971, S. Cook showed that not only the Traveling Salesman Problem, but also all other problems in the class *NP*, can be reduced to a satisfiability problem in the propositional calculus. For that reason the satisfiability problem for propositional logic is called *NP-complete.*

Exercise 2.31. [Keisler; appearance in S.C. Kleene [14], p. 67] Brown, Jones and Smith are suspected of income tax evasion. They testify under oath as follows.
BROWN: Jones is guilty and Smith is innocent.

JONES: If Brown is guilty, then so is Smith.
SMITH: I'm innocent, but at least one of the others is guilty.
Let B, J, S be the statements 'Brown is innocent', 'Jones is innocent', 'Smith is innocent', respectively. Express the testimony of each suspect by a formula in our logical symbolism, and write out the truth tables for these three formulas (in parallel columns). Now answer the following questions.
a) Are the testimonies of the three suspects consistent, i.e., is the conjunction of these testimonies consistent?
b) The testimony of one of the suspects follows from that of another. Which from which?
c) Assuming everybody is innocent, who committed perjury?
d) Assuming everyone's testimony is true, who is innocent and who is guilty?
e) Assuming that the innocent told the truth and the guilty told lies, who is innocent and who is guilty?

Exercise 2.32. [W. Ophelders] The football clubs Pro, Quick and Runners play a football tournament. The trainers of these clubs make the following statements.
Trainer of Pro: If the Runners win the tournament, then Quick does not.
Trainer of Quick: We or the Runners win the tournament.
Trainer of the Runners: We win the tournament.
Express the three statements by formulas in our logical symbolism and write out the truth tables for these three formulas. Next answer the following questions, supposing there can be at most one winner.
a) Assuming everyone's statement is true, which club wins the tournament?
b) Assuming only the trainer of the winning club makes a true statement, which club wins the tournament?

Exercise 2.33. Find formulas composed from P, Q, R, $\wedge$, $\vee$ and $\neg$ only, whose truth tables have the following value columns:

P	Q	R	(a)	(b)	(c)	(d)
1	1	1	0	1	0	0
1	1	0	0	0	0	1
1	0	1	0	0	0	1
1	0	0	0	0	0	1
0	1	1	1	1	0	1
0	1	0	0	0	0	0
0	0	1	0	1	0	1
0	0	0	0	0	0	1

Exercise 2.34. Let $A \downarrow B$ be defined by the following truth table:

A	B	$A \downarrow B$
1	1	0
1	0	1
0	1	1
0	0	1

$A \downarrow B$ may be read as 'not A or not B'. Prove that $\neg$, $\vee$ and $\wedge$, and hence each of the 16 binary truthfunctional connectives, can be expressed in terms of $\downarrow$.

Exercise 2.35. A set of binary truthfunctional connectives is *independent* iff none of the members of the set can be expressed in terms of the other members of the set.
i) Show that $\{\wedge, \vee, \neg\}$ is not independent.
ii) Show that $\{\wedge, \neg\}$, $\{\vee, \neg\}$ and $\{\rightarrow, \neg\}$ are independent and complete sets of truthfunctional binary connectives.

Exercise 2.36. Show that there are only two binary connectives, namely, $\uparrow$ (the Sheffer stroke) and $\downarrow$ (see Exercise 2.34) such that every binary truthfunctional connective can be expressed in it.

Exercise 2.37. Construct formulas in conjunctive normal form which have the same truth table as the following formulas:
i) $(P \rightarrow (Q \rightarrow P)) \wedge (P \rightarrow Q \vee P)$
ii) $(P \rightarrow \neg(Q \rightarrow P)) \wedge (P \rightarrow Q \wedge P)$
iii) $(P \rightarrow \neg(Q \rightarrow P)) \vee (P \rightarrow Q \wedge P)$

2.6 Syntax: Provability and Deducibility

By now it will be clear that there are a great many, in fact even infinitely many, valid formulas. And given premisses $A_1, \ldots, A_n$, there are infinitely many valid consequences of those premisses. The question now arises whether it is possible to select a few valid formulas, to be called *logical axioms*, together with certain rules – which applied to valid formulas produce (or generate) new valid formulas – such that any valid formula can be obtained (or generated) by a finite number of applications of the given rules to the selected logical axioms. This question can be answered positively, which means that in a certain sense we have reduced the big collection of valid formulas to a surveyable subset: any formula in the big collection of valid formulas can be generated by the given rules from formulas in the subset.

There are several possibilities for choosing the logical axioms and rules such that the desired goal is accomplished. In this section one of them is presented, namely, a system for propositional logic developed by Frege, and adapted by Russell and Hilbert. Henceforth, we shall speak of a *Hilbert-type* system. In Section 2.8 two other, more recent, systems will be treated which achieve the same goal.

One may design production methods satisfying the following two conditions:

(I) the production method produces in the course of time only formulas which are valid, and, more generally,

(II) the production method if applied to certain formulas given as premisses, only produces formulas which are a valid consequence of those premisses.

There are in fact many such production methods, each of them consisting of (i) a set of valid formulas, and (ii) a set of rules of inference. One such production method satisfying (I) and (II) can be obtained by taking:

(i) All formulas of any of the forms $A \to (B \to A)$ and
$$(A \to B) \to ((A \to (B \to C)) \to (A \to C));$$
We have seen in Theorem 2.7, 1 and 2, that such formulas are valid. We call these formulas (logical) *axioms for the connective* $\to$.

(ii) As the sole rule of inference, called the $\to$*-rule* or *Modus Ponens* (*MP*), we take the operation of passing from two formulas of the respective forms D and $D \to E$ to the formula E, for any choice of formulas D and E.

$$\text{Modus Ponens } (MP): \quad \frac{D \qquad\qquad D \to E}{E}$$

In an inference by this rule, the formulas D and $D \to E$ are the *premisses*, and E is the *conclusion*. The following statements can easily be checked:

(α) Any interpretation that makes the premisses of the rule true, also makes the conclusion of the rule true. For our particular rule MP: for any interpretation i, if $i(D) = 1$ and $i(D \to E) = 1$, then $i(E) = 1$, and consequently

(β) If all premisses of the rule are valid, then also the conclusion of the rule is valid. For our particular rule MP: if $\models D$ and $\models D \to E$, then $\models E$ (Theorem 2.11). Our rule of inference may be applied zero, one, two or more times to formulas of the form mentioned in (i) or to formulas which we have already generated earlier.

Example 2.12. This production method yields, among other things, the following formulas for any choice of the formula A:

1. $A \to (A \to A)$ This is a formula of the form $A \to (B \to A)$, taking $B = A$.
2. $(A \to (A \to A)) \to ((A \to ((A \to A) \to A)) \to (A \to A))$ This is a formula of the form $(A \to B) \to ((A \to (B \to C)) \to (A \to C))$, taking $B = A \to A$ and $C = A$.
3. $(A \to ((A \to A) \to A)) \to (A \to A)$ This formula is obtained by an application of Modus Ponens to 1 and 2.
4. $A \to ((A \to A) \to A)$ This formula is of the form $A \to (B \to A)$, taking $B = A \to A$.
5. $A \to A$ This formula is obtained by an application of Modus Ponens to 3 and 4.

Schematically:

$$\cfrac{A \to ((A \to A) \to A) \qquad \cfrac{A \to (A \to A) \quad (A \to (A \to A)) \to ((A \to ((A \to A) \to A)) \to (A \to A))}{(A \to ((A \to A) \to A)) \to (A \to A)}\, MP}{A \to A}\, MP$$

This schema is called a (logical, Hilbert-type) *proof* of the formula $A \to A$ and $A \to A$ is called (logically) *provable*, because there exists such a schema using only logical axioms and Modus Ponens. Note that each of the formulas in this schema, and $A \to A$ in particular, is produced by our production method and that each of these formulas is valid, since we started with valid formulas and since Modus Ponens applied to valid formulas only yields formulas which are valid (Theorem 2.11 or (β) above).

Example 2.13. The production method described above applied to the formulas $A \to B$ and $B \to C$, for instance, yields the following formulas:

1. $A \to B$ This formula is a given premiss.
2. $(A \to B) \to ((A \to (B \to C)) \to (A \to C))$ This formula is of the appropriate form.
3. $(A \to (B \to C)) \to (A \to C)$ Obtained by applying Modus Ponens to 1 and 2.
4. $B \to C$ This formula is a given premiss.
5. $(B \to C) \to (A \to (B \to C))$ This is a formula of the form $A \to (B \to A)$, taking $A = B \to C$ and $B = A$.
6. $A \to (B \to C)$ This formula is obtained by applying Modus Ponens to 4 and 5.
7. $A \to C$ This formula is obtained by an application of Modus Ponens to 6 and 3.

Schematically:

$$\dfrac{\dfrac{\overset{\text{premiss}}{B \to C} \quad \overset{\text{axiom 1}}{(B \to C) \to (A \to (B \to C))}}{A \to (B \to C)} \qquad \dfrac{\overset{\text{premiss}}{A \to B} \quad \overset{\text{axiom 2}}{(A \to B) \to ((A \to (B \to C)) \to (A \to C))}}{(A \to (B \to C)) \to (A \to C)}}{A \to C}$$

This schema is called a (logical, Hilbert-type) *deduction* of $A \to C$ from the premisses $A \to B$ and $B \to C$ and $A \to C$ is said to be *deducible* from the premisses $A \to B$ and $B \to C$, using only these premisses, logical axioms and Modus Ponens. Note that each of the formulas in this schema, and $A \to C$ in particular, is produced by our production method applied to the premisses $A \to B$ and $B \to C$, and that each of these formulas is a valid consequence of the premisses $A \to B$ and $B \to C$, since we started with valid formulas, the premisses $A \to B$ and $B \to C$ only, and because of (α) above.

It will be clear now that any production method, consisting of (i) a set of valid formulas and (ii) a set of rules of inference satisfying (α) and (β), will satisfy the conditions (I) and (II), mentioned in the beginning of this section.

One can prove (see Exercise 2.44) that Peirce's law, $((A \to B) \to A) \to A$, although it contains only the connective $\to$, is not generated by the production method consisting of the two logical axioms for $\to$ and Modus Ponens. This raises the question whether there is a *complete* production method satisfying I and II, i.e., a production method which in the course of time generates *all* valid formulas and, more generally, which generates, if applied to certain formulas, given as premisses, *all* valid consequences of those premisses. The answer to this question is affirmative. In Section 2.9 we shall prove that the production method consisting of all formulas of any of the forms shown after the symbol $\models$ in Theorem 2.7, and of the sole rule of inference, Modus Ponens, is complete. For convenience these formulas are again listed below and will be called (logical) *axioms for (classical) propositional logic*.

1. $A \to (B \to A)$
2. $(A \to B) \to ((A \to (B \to C)) \to (A \to C))$
3. $A \to (B \to A \wedge B)$

4a $A \wedge B \to A$

4b $A \wedge B \to B$
5a $A \to A \vee B$
5b $B \to A \vee B$
6. $(A \to C) \to ((B \to C) \to (A \vee B \to C))$
7. $(A \to B) \to ((A \to \neg B) \to \neg A)$
8. $\neg\neg A \to A$
9. $(A \to B) \to ((B \to A) \to (A \rightleftarrows B))$
10a $(A \rightleftarrows B) \to (A \to B)$
10b $(A \rightleftarrows B) \to (B \to A)$

Numbers 1 and 2 concern axioms for the connective $\to$, numbers 3 and 4 concern axioms for $\wedge$, numbers 5 and 6 concern axioms for $\vee$, numbers 7 and 8 concern axioms for $\neg$ and numbers 9 and 10 concern axioms for $\rightleftarrows$. Notice that in a sense they describe the typical properties of the connective in question; for instance, the axioms for $\wedge$ will not hold if we replace $\wedge$ by $\vee$.

These forms themselves will be called *axiom schemata*. Each schema includes infinitely many axioms, one for each choice of the formulas denoted by A, B, C. For example, corresponding to 1 in Theorem 2.7, we have as Axiom Schema 1: $A \to (B \to A)$. Particular axioms in this schema are $P \to (P \to P)$, $P \to (Q \to P)$, $Q \to (P \to Q)$, $\neg P \to (Q \wedge R \to \neg P)$, $(P \to (\neg Q \to P)) \to (R \to (P \to (\neg Q \to P)))$, etc.

The choice of the logical axioms is a subtle matter. For instance, if one would replace axiom schema 8, $\neg\neg A \to A$, by its converse, $A \to \neg\neg A$, then the resulting system would not be complete, in particular, the resulting system would not be able to generate Peirce's law, $((A \to B) \to A) \to A$. Also, if one replaces axiom 8, $\neg\neg A \to A$, by $\neg A \to (A \to B)$ one obtains intuitionistic propositional logic, which is completely different from classical logic; see Chapter 8. Small changes may have far reaching consequences!

Example 2.14. For illustration, let us show that from the premisses
$P \to W$: I will pay them for fixing our TV [P] only if it works [W].
$\neg W$: Our TV still does not work.
the logical consequence $\neg P$ (I will not pay) can be generated by using the logical axioms 1 and 7 and by three applications of Modus Ponens.

$$\dfrac{\dfrac{\overset{\text{prem}}{\neg W} \qquad \overset{\text{axiom 1}}{\neg W \to (P \to \neg W)}}{P \to \neg W} \qquad \dfrac{\overset{\text{prem}}{P \to W} \qquad \overset{\text{axiom 7}}{(P \to W) \to ((P \to \neg W) \to \neg P)}}{(P \to \neg W) \to \neg P}}{\neg P}$$

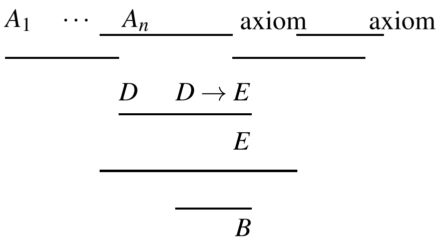
The schema above is called a (logical, Hilbert-type) *deduction* of $\neg P$ from the premisses $P \to W$ and $\neg W$ and we say that $\neg P$ is (logically) *deducible from* $P \to W$ and $\neg W$, meaning that there exists a (logical, Hilbert-type) deduction of $\neg P$ from $P \to W$ and $\neg W$.

Definition 2.11 (Deduction; Deducible). Let $B, A_1, \ldots, A_n$ be formulas.

1. A (logical, Hilbert-type) *deduction* of B *from* $A_1,\ldots,A_n$ (in classical propositional logic) is a finite list $B_1,\ldots,B_k$ of formulas, such that
   (a) $B = B_k$ is the last formula in the list, and
   (b) each formula in the list is either one of $A_1,\ldots,A_n$, or one of the axioms of propositional logic (see Theorem 2.7), or is obtained by an application of Modus Ponens to a pair of formulas preceding it in the list.
2. B is *deducible from* $A_1,\ldots,A_n$:= there exists a (logical, Hilbert-type) deduction of B from $A_1,\ldots,A_n$.
   Notation: $A_1,\ldots,A_n \vdash B$, where the symbol $\vdash$ may be read 'yields'. If there does not exist a deduction of B from $A_1,\ldots,A_n$ this is written as $A_1,\ldots,A_n \nvdash B$ as shorthand for: not $A_1,\ldots,A_n \vdash B$.
3. In case $n = 0$, i.e., in case there are no premisses, these definitions reduce to:
   A (logical, Hilbert-type) *proof* of B is a finite list of formulas with B as last formula in the list, such that every formula in the list is either an axiom of propositional logic or obtained by Modus Ponens to formulas earlier in the list.
   B is (logically) *provable* := there exists a (logical, Hilbert-type) proof of B.
   Notation: $\vdash B$
4. For Γ a (possibly infinite) set of formulas, B is *deducible from* Γ, if there is a finite list $A_1,\ldots,A_n$ of formulas in Γ such that $A_1,\ldots,A_n \vdash B$.
   Notation: $\Gamma \vdash B$.

Example 2.15. We have seen in Example 2.13 that $A \to B,\ B \to C \vdash A \to C$ and in Example 2.14 that $P \to W, \neg W \vdash \neg P$. And also in Example 2.12 that $\vdash A \to A$.

So, $A_1,\ldots,A_n \vdash B$, in words: B is *deducible from* $A_1,\ldots,A_n$, if and only if there exists a finite schema of the form



And in case there are no premisses $A_1,\ldots,A_n$, i.e., $n = 0$, we say that $\vdash B$, in words: B is (logically) *provable* or *deducible*.

Example 2.16. Consider the following sequence of formulas:

$$\dfrac{\dfrac{\overset{\text{premiss}}{A \wedge B} \qquad \overset{\text{4b}}{A \wedge B \to B}}{B}\,MP \qquad \dfrac{\dfrac{\overset{\text{premiss}}{A \wedge B} \qquad \overset{\text{4a}}{A \wedge B \to A}}{A}\,MP \qquad \overset{\text{premiss}}{A \to (B \to C)}}{B \to C}\,MP}{C}\,MP$$

For each choice of formulas $A,\ B,\ C$, this sequence of formulas is a deduction of C from $A \rightarrow (B \rightarrow C)$ and $A \wedge B$. Hence, C is deducible from $A \rightarrow (B \rightarrow C)$ and $A \wedge B$; i.e., $A \rightarrow (B \rightarrow C),\ A \wedge B \vdash C$.

The notion of logical consequence, $A_1, \ldots, A_n \models B$, is in terms of the truth or falsity and hence in terms of the *meaning* of the formulas involved. Therefore, this notion of logical consequence is a semantic notion. But the notion of (logical) deducibility, $A_1, \ldots, A_n \vdash B$, is in terms of the *forms* of the formulas involved. One does not have to know the meaning of the connectives, one only has to distinguish the form of the formulas involved. Therefore, this notion is a syntactic notion.

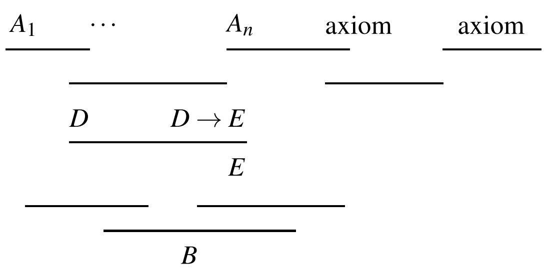
In $A_1, \ldots, A_n \vdash B$ one may think of the premisses $A_1, \ldots, A_n$ as being the (non-logical) axioms of Euclid ($\pm$ 300 B.C.) for geometry, the axioms of Peano for arithmetic (see Chapter 5), the axioms of Zermelo - Fraenkel for set theory (see Chapter 3) or the laws of Newton for classical mechanics.

The premisses $A_1, \ldots, A_n$, formulated in an appropriate formal language, constitute what one calls a (formal) *theory*: Euclid's geometry, Peano's arithmetic, the set theory of Zermelo - Fraenkel, Newton's mechanics, and so on. Each science is continually trying to re-adjust its foundations, as formulated in its premisses. For instance, Cantor's naive set theory had to be replaced by the set theory of Zermelo - Fraenkel (see Chapter 3) and Newton's (classical) mechanics by Einstein's theory of relativity.

Of course, we want that our production method, consisting of the (logical) axioms for propositional logic and Modus Ponens, is *sound*, that is, when applied to given premisses $A_1, \ldots, A_n$, it should generate only formulas which are a logical (or valid) consequence of $A_1, \ldots, A_n$. This is indeed the case, as stated in the following *soundness theorem*.

Theorem 2.20 (Soundness theorem).
(a): If $A_1, \ldots, A_n \vdash B$, then $A_1, \ldots, A_n \models B$, or, equivalently,
(a') if $A_1, \ldots, A_n \not\models B$, then $A_1, \ldots, A_n \nvdash B$.
(b): In case $n = 0$, i.e., there are no premisses: if $\vdash B$, then $\models B$.
(c): If $\Gamma \vdash B$, then $\Gamma \models B$.

Proof. Suppose $A_1, \ldots, A_n \vdash B$, i.e., there is a finite schema of the form



Note the following:
i) Each axiom of propositional logic has the value 1 in each line of the truth table.

ii) For all lines of the truth table, given that in an application of Modus Ponens the premisses D and $D \to E$ have the value 1, the conclusion E has the value 1 as well.

We have to show that $A_1, \ldots, A_n \models B$. So, suppose that the premisses $A_1, \ldots, A_n$ are 1 in a given line of the truth table. Then it follows from i) and ii) that, going from top to bottom in the deduction of B from $A_1, \ldots, A_n$, every formula in the deduction has value 1 in the given line. Hence, in particular, B has value 1 in that same line of the truth table. □

One may illustrate this proof by a concrete example, for instance, for the case that $A \to (B \to C),\ A \wedge B \vdash C$.

Corollary 2.2 (Simple consistency).
There is no formula B such that both $\vdash B$ *and* $\vdash \neg B$.

Proof. Suppose $\vdash B$ and $\vdash \neg B$ for some B. Then according to the soundness theorem 2.20, $\models B$ and $\models \neg B$. Contradiction. □

We hope that the production method, consisting of the (logical) axioms for propositional logic and Modus Ponens, is *complete*, that is, that every valid consequence of given premisses $A_1, \ldots, A_n$ may be logically) deduced from these premisses. This is indeed the case, as is stated in the following theorem, which will be proved in Section 2.9 and in Exercise 2.59.

Theorem 2.21 (Completeness theorem).
(a): If $A_1, \ldots, A_n \models B$*, then* $A_1, \ldots, A_n \vdash B$*, or, equivalently,*
(a') if $A_1, \ldots, A_n \not\vdash B$*, then* $A_1, \ldots, A_n \not\models B$*.*
(b): In case $n = 0$*, i.e., there are no premisses: if* $\models B$*, then* $\vdash B$*.*
(c): If $\Gamma \models B$*, then* $\Gamma \vdash B$*.*

By the *soundness* of the axiomatic-deductive system for (classical) propositional logic we mean that *at most* certain formulas are provable, namely only those which are valid; by the *completeness* we mean that *at least* certain formulas are provable, namely, all which are valid. By the end of Section 2.9 we shall have proved the completeness theorem and hence (combining completeness and soundness) have shown the following equivalences:

$$\begin{array}{rcl} A_1, \ldots, A_n \models B & \text{iff} & A_1, \ldots, A_n \vdash B \\ \Gamma \models B & \text{iff} & \Gamma \vdash B \\ \models B & \text{iff} & \vdash B \end{array}$$

There are a number of arguments underscoring the *philosophical meaning of the completeness theorem*, which justify taking the trouble to prove this theorem.

1. The completeness theorem tells us that any correct argument (in the object language) has a rational reconstruction which has the standard form described in the definition of $A_1, \ldots, A_n \vdash B$. Arguments in science and in daily life usually do not proceed in the way described in the definition of $A_1, \ldots, A_n \vdash B$, but according to the completeness theorem for any such correct argument there is a rational reconstruction which does.

2. Note that whether B is deducible from $A_1, \ldots, A_n$ or not only depends on the *form* of the formulas $A_1, \ldots, A_n$ and B. Hence, the question whether B is a valid consequence of $A_1, \ldots, A_n$ or not has been reduced to a question about the form of the formulas $A_1, \ldots, A_n$ and B.
3. We have defined the intuitive notion of 'B is a logical consequence of $A_1, \ldots, A_n$' in two completely different ways; we have given a semantic definition in terms of truth values ($A_1, \ldots, A_n \models B$) and a syntactic one in terms of logical axioms and the rule Modus Ponens ($A_1, \ldots, A_n \vdash B$). That these two notions turn out to be equivalent suggests that our definitions indeed capture the corresponding intuitive notion.
4. We have given a mathematically precise definition of the intuitive notion of logical consequence in order to make this notion mathematically manageable, which is necessary if one wants to prove in a precise way certain statements about this notion. Now it is safe to assume that

   a) if B is intuitively a logical consequence of $A_1, \ldots, A_n$, then $A_1, \ldots, A_n \models B$.
   According to the completeness theorem,
   b) if $A_1, \ldots, A_n \models B$, then $A_1, \ldots, A_n \vdash B$.
   An analysis of the axioms and rules of propositional logic indicates that
   c) if $A_1, \ldots, A_n \vdash B$, then B is intuitively a logical consequence of $A_1, \ldots, A_n$.

   (a), (b) and (c) show that the intuitive notion of logical consequence and the mathematical notions of $A_1, \ldots, A_n \models B$ and of $A_1, \ldots, A_n \vdash B$ coincide extensionally.
5. In Chapter 4 we shall extend the notion of valid or logical consequence and of (logical) deducibility to (classical) *predicate logic*. Then we shall prove that these notions are again equivalent (soundness and completeness). On that occasion we shall further elaborate on the meaning of the completeness theorem in the case of predicate logic.

In Example 2.14 we have constructed a logical deduction of $\neg P$ from the premisses $P \to W$ and $\neg W$, hence, $P \to W, \neg W \vdash \neg P$, where P and W were atomic formulas. More generally, in the same way one can show that for arbitrary formulas A and B, $A \to B, \neg B \vdash \neg A$. That is, the rule *Modus Tollens*

$$\frac{A \to B \qquad \neg B}{\neg A}$$

is a derived rule, that from now on may be used in the construction of (logical) deductions. There are many more derived rules, for instance, see Exercise 2.39.

Exercise 2.38. Translate the following arguments in logical terminology and check whether the (putative) conclusion is *deducible* from the premisses. If so, give a deduction, using the logical axioms $K \to (R \to K)$ and $(R \to K) \to ((R \to \neg K) \to \neg R)$. If not, then why not?
a) If it rains [R], then John will not come [$\neg C$]. John will come. Therefore: it does not rain.
b) Only if it rains [R], John will not come [$\neg C$]. John will come. Therefore: it does not rain.

Exercise 2.39. By constructing appropriate deductions, show that

(a)	$A,\, A \to B \vdash B$	(f)	$B \vdash A \vee B$
(b)	$A,\, B \vdash A \wedge B$	(g)	$\neg\neg A \vdash A$
(c)	$A \wedge B \vdash A$	(h)	$A \to B,\, B \to A \vdash A \rightleftarrows B$
(d)	$A \wedge B \vdash B$	(i)	$A \rightleftarrows B \vdash A \to B$
(e)	$A \vdash A \vee B$	(j)	$A \rightleftarrows B \vdash B \to A$

Hence, from now on, the following derived rules may be used in the construction of (logical) deductions:

$$\frac{A \quad B}{A \wedge B} \qquad \frac{A \wedge B}{A} \qquad \frac{A \wedge B}{B} \qquad \frac{A}{A \vee B} \qquad \frac{B}{A \vee B} \qquad \frac{\neg\neg A}{A}$$

Exercise 2.40. Prove that $A, \neg A \vdash B$ by using the following axioms:
axiom 1 (a): $A \to (\neg B \to A)$ axiom 1 (b): $\neg A \to (\neg B \to \neg A)$
axiom 7: $(\neg B \to A) \to ((\neg B \to \neg A) \to \neg\neg B)$ axiom 8: $\neg\neg B \to B$

Exercise 2.41. By using the soundness theorem show that

(a) not $P \vee Q \vdash P \wedge Q$, (c) not $P \vdash Q$,
(b) not $P \to Q \vdash Q \to P$, (d) not $P \to Q \vdash P \wedge Q$.

Note that in order to show that $A \vdash B$, it suffices to exhibit at least one logical deduction of B from A; but in order to show that not $A \vdash B$, one has to prove that no logical deduction of B from A can exist, in other words, that any deduction is not a deduction of B from A. In order to prove the latter, it suffices – according to the soundness theorem – to show that $A \not\models B$.

Exercise 2.42. Prove or refute: $P \to Q,\ P \vdash R \vee Q$ either by giving a deduction of $R \vee Q$ from $P \to Q$ en P, using the logical axiom $B \to A \vee B$, or by showing that such a deduction cannot exist.

Exercise 2.43. Translate the following argument in logical terminology and check whether the (putative) conclusion is *deducible* from the premisses. If so, give a deduction, using the logical axioms $A \to (B \to A)$, $(A \to B) \to ((A \to \neg B) \to \neg A)$ and $\neg\neg A \to A$. If not, why not?
If John succeeds [S], then John works hard [H].
If John is not intelligent [$\neg I$], then John does not succeed.
Therefore: if John is intelligent, then John works hard.

Exercise 2.44. Consider a system of three truth values, 0, 1 and 2, of which 0 is the only designated truth value, and let the truth table of $\to$ be as follows.

A	B	$A \to B$
0	0	0
0	1	1
0	2	2
1	0	0
1	1	0
1	2	2
2	0	0
2	1	0
2	2	0

Show that for any choice of formulas A, B, C

a) for every interpretation i, $i(A \to (B \to A)) = 0$,
b) for every interpretation i, $i((A \to B) \to ((A \to (B \to C)) \to (A \to C))) = 0$,
c) for every interpretation i, if $i(A) = 0$ and $i(A \to B) = 0$, then $i(B) = 0$,
d) for some interpretation i, $i(((A \to B) \to A) \to A) \neq 0$.

Conclude that Peirce's law, $((A \to B) \to A) \to A$, is independent of $A \to (B \to A)$ and $(A \to B) \to ((A \to (B \to C)) \to (A \to C))$, in other words, that Peirce's law is not generated by the production method consisting of only the two axioms for $\to$ and Modus Ponens.

2.7 Syntax: Meta-logical Results

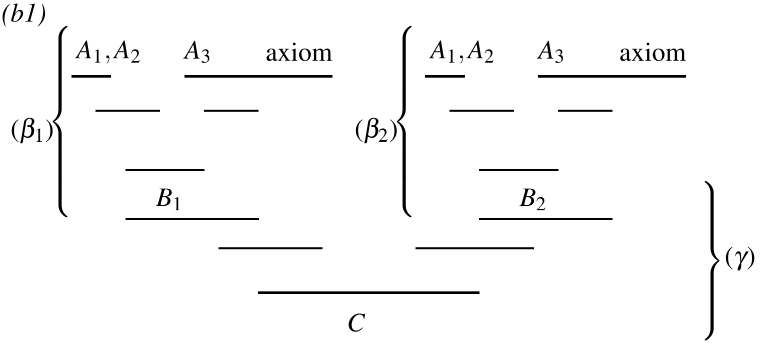
In this section (logical) proofs and deductions in the object-language will be studied, using (of necessity) informal proofs and deductions in the meta-language. The main results are the Deduction theorem, and the Introduction and Elimination rules. Given premisses $A_1, \ldots, A_n$ and given a formula B, these theorems are crucial in facilitating the search for a logical deduction of B from $A_1, \ldots, A_n$, if there is one. Next Gentzen's system of Natural Deduction is presented. It is shown that any formula which is logically provable in this system is also provable in the proof-system of Section 2.6, and conversely.

In Section 2.6 we defined a (logical) deduction of B from premisses $A_1, \ldots, A_n$ as being a finite sequence of formulas which satisfies certain conditions. It is important to realize that whether a given sequence of formulas is a (logical) deduction or not only depends on the *form* of the formulas in the sequence. In other words, whether a given sequence of formulas is a (logical) deduction can be checked mechanically; one can write a computer program to check the correctness of a given putative (logical) deduction. An example is *Automath*, developed by N.G. de Bruijn [3] and others at Eindhoven University.

It is also important to distinguish between logical deductions (of formulas) in the object language and informal proofs of certain statements *about* logical deductions. For instance, in Theorem 2.22 (b1) we will prove informally that if $A_1, A_2, A_3 \vdash B_1$ and $A_1, A_2, A_3 \vdash B_2$ and $B_1, B_2 \vdash C$, then $A_1, A_2, A_3 \vdash C$. This theorem is *about* logical proofs and deductions in the object-language; however, the formulation and the (informal) proof of this theorem are given in the meta-language. Notice that this Theorem is the syntactic counterpart of Exercise 2.26.

Theorem 2.22.
(a1) $A_1, A_2, A_3 \vdash A_1$, $A_1, A_2, A_3 \vdash A_2$, $A_1, A_2, A_3 \vdash A_3$.
(a2) More generally: $A_1, \ldots, A_i, \ldots, A_n \vdash A_i$ *for* $i = 1, \ldots, n$.
(b1) If $A_1, A_2, A_3 \vdash B_1$ *and* $A_1, A_2, A_3 \vdash B_2$ *and* $B_1, B_2 \vdash C$, *then* $A_1, A_2, A_3 \vdash C$.
(b2) More generally, for any $n, k \geq 0$*: if* $A_1, \ldots, A_n \vdash B_1$ *and* $\ldots$ *and* $A_1, \ldots, A_n \vdash B_k$ *and* $B_1, \ldots, B_k \vdash C$, *then* $A_1, \ldots, A_n \vdash C$.

Proof. (a1) For each $i, 1 \leq i \leq 3$, A_i itself is a (logical) deduction of A_i from A_1, A_2, A_3. In the definition of a logical deduction it is not required that all the premisses are actually used; they *may* be used, but not necessarily so.
(a2) is shown similarly.
(b1)



Assume $A_1, A_2, A_3 \vdash B_1$, $A_1, A_2, A_3 \vdash B_2$ and $B_1, B_2 \vdash C$. That is, there are deductions (β_1) and (β_2) of B_1 and B_2 respectively, from A_1, A_2, A_3 and there is a deduction (γ) of C from B_1, B_2. By replacing the premisses B_1 and B_2 in (γ) by the deductions (β_1) and (β_2), we obtain a (logical) deduction of C from A_1, A_2, A_3. Hence, $A_1, A_2, A_3 \vdash C$.
(b2) is shown similarly. □

If we take in Theorem 2.22 (b1) $B_1 = B_2 = A_3 = A$, we obtain the following result.

Corollary 2.3. *If $A \vdash C$, then A_1, A_2, $A \vdash C$.*
More generally: If $A \vdash C$, then $A_1, \ldots, A_{n-1}$, $A \vdash C$.

Proof. In the definition of $A_1, \ldots, A_{n-1}$, $A \vdash C$ it is not required that each of the assumption formulas $A_1, \ldots, A_{n-1}$ actually occur in the deduction. □

Theorem 2.22 can be reformulated in set-theoretic terms: let $L(A_1, \ldots, A_n)$, called the *logic of* $A_1, \ldots, A_n$, be the set of all formulas that are deducible from $A_1, \ldots, A_n$. Then Theorem 2.22 says that i) for each $i, 1 \leq i \leq n$, A_i is in $L(A_1, \ldots, A_n)$, and ii) if each of $B_1, \ldots, B_k$ is in $L(A_1, \ldots, A_n)$ and $B_1, \ldots, B_k \vdash C$, then C is in $L(A_1, \ldots, A_n)$.

Since in Corollary 2.3 the premisses $A_1, \ldots, A_{n-1}$ are not relevant to C, Corollary 2.3, which just has been shown for classical logic, does not hold for the so-called *relevance logic*; see Section 6.10.

Let us consider the following four expressions:
(i) $\models A \rightarrow B$ i.e., $A \rightarrow B$ is valid,
(ii) $A \models B$ i.e., B is a valid consequence of A,
(iii) $\vdash A \rightarrow B$ i.e., $A \rightarrow B$ is (logically) provable,
(iv) $A \vdash B$ i.e., B is (logically) deducible from A.
(i) and (ii) are *semantic notions*, i.e., they are concerned with the *meaning* of the formulas in question; (iii) and (iv) are *syntactic notions*, i.e., they are concerned with the *form* of the formulas in question.

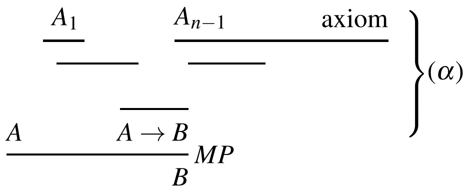
In Theorem 2.4 we have already shown that (i) and (ii) are equivalent. In Theorems 2.23 and 2.24 we will prove that (iii) and (iv) are equivalent.

In the soundness theorem (Theorem 2.20) we have shown that (iii) implies (i) and that (iv) implies (ii). The converses of these results, (i) implies (iii) and (ii) implies (iv), will be shown in Section 2.9.

So, by the end of Section 2.9 we shall have proved that (i), (ii), (iii) and (iv) are equivalent. But remember that 'if $\models A$, then $\models B$' is a weaker statement than (ii), $A \models B$ (see Theorem 2.11). Consequently, 'if $\vdash A$, then $\vdash B$' is a weaker statement than (iv), $A \vdash B$.

Theorem 2.23. *(a) If* $\vdash A \to B$*, then* $A \vdash B$*.* *(b) More generally, for any* $n \geq 1$*, if* $A_1, \ldots, A_{n-1} \vdash A \to B$*, then* $A_1, \ldots, A_{n-1},\ A \vdash B$*.*

Proof. (b) Suppose $A_1, \ldots, A_{n-1} \vdash A \to B$, i.e., there is a deduction (α) of $A \to B$ from $A_1, \ldots, A_{n-1}$.



By adding one more premiss, A, to this deduction and one more application of Modus Ponens, one obtains a deduction of B from $A_1, \ldots, A_{n-1},\ A$. □

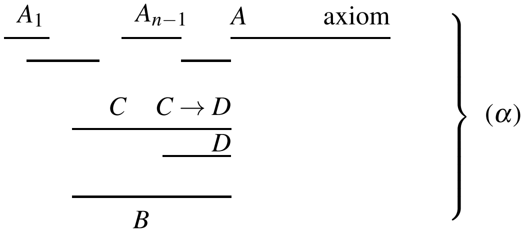
2.7.1 Deduction Theorem; Introduction and Elimination Rules

In order to establish an implication 'if A, then B', one often assumes A and then continues to conclude B. The following theorem, called the deduction theorem, which is the converse of Theorem 2.23, captures this idea in a precise form: in order to establish that $A_1, \ldots, A_{n-1} \vdash A \to B$, it suffices to show that $A_1, \ldots, A_{n-1}, A \vdash B$.

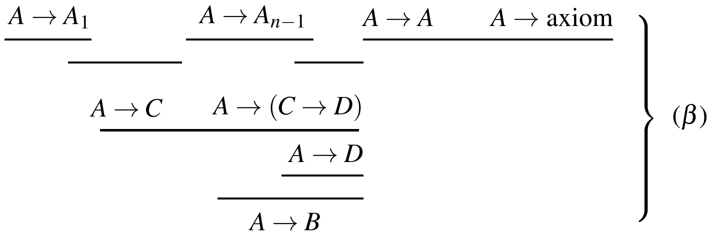
That the deduction theorem is a very useful tool may be seen from the following. In order to show that $\vdash A \to ((A \to B) \to B)$, it suffices by the deduction theorem to show that $A \vdash (A \to B) \to B$. Likewise, in order to show the latter statement it suffices to prove $A,\ A \to B \vdash B$; and this is very easy (one application of Modus Ponens suffices), while to show that $\vdash A \to ((A \to B) \to B)$ directly is much more complicated.

Theorem 2.24 (Deduction theorem, Herbrand 1930).
(a) If $A \vdash B$*, then* $\vdash A \to B$*. More generally,*
(b) If $A_1, \ldots, A_{n-1},\ A \vdash B$*, then* $A_1, \ldots, A_{n-1} \vdash A \to B$*.*

Proof. (b) Suppose $A_1, \ldots, A_{n-1},\ A \vdash B$, i.e., there is a (logical) deduction (α) of B from the premisses $A_1, \ldots, A_{n-1},\ A$. Below we shall change (α) step by step into a (logical) deduction (γ) of $A \to B$ from $A_1, \ldots, A_{n-1}$, hence showing that $A_1, \ldots, A_{n-1} \vdash A \to B$.



The first step consists in prefixing the symbols $A \to$ to each formula occurring in (α). This results in the schema (β).



Although the last formula in (β) is $A \to B$, (β) itself is not a deduction of $A \to B$ from $A_1, \ldots, A_{n-1}$ for the following reasons:
(i) (β) does not start with logical axioms or premisses $A_1, \ldots, A_{n-1}$, and (ii)

$$\frac{A \to C \qquad A \to (C \to D)}{A \to D}$$

is not an application of Modus Ponens.

However, by inserting appropriate formulas into (β), one can transform (β) into a (logical) deduction (γ) of $A \to B$ from $A_1, \ldots, A_{n-1}$ as follows.

1. For $1 \le j \le n-1$ replace $A \to A_j$ at the top in (β) by the following:

$$\frac{A_j \qquad \overset{\text{axiom 1}}{A_j \to (A \to A_j)}}{A \to A_j}\ MP$$

2. Replace $A \to A$ at the top in (β) by the (logical) proof of $A \to A$, given in Section 2.6.
3. Replace $A \to$ axiom at the top in (β) by the following:

$$\frac{\text{axiom} \qquad \overset{\text{axiom 1}}{\text{axiom} \to (A \to \text{axiom})}}{A \to \text{axiom}}\ MP$$

4. Replace

$$\frac{A \to C \qquad A \to (C \to D)}{A \to D}$$

by the following:

$$\dfrac{\dfrac{A \to C \qquad \overset{\text{axiom 2}}{(A \to C) \to ((A \to (C \to D)) \to (A \to D))}}{(A \to (C \to D)) \to (A \to D)}\,MP \qquad A \to (C \to D)}{A \to D}$$

Each formula of the resulting sequence (γ) either is one of $A_1, \dots, A_{n-1}$ or is a logical axiom or comes from two preceding formulas in the sequence by Modus Ponens, and the last formula of the sequence is $A \to B$. So (γ) is a deduction of $A \to B$ from $A_1, \dots, A_{n-1}$. □

In Exercise 2.58 the proof of the deduction theorem is applied to a deduction of $Q \vee R$ from $P \to Q$ and P in order to obtain a deduction of $P \to Q \vee R$ from $P \to Q$.

Example 2.17. In Example 2.16 we have seen that $A \to (B \to C),\ A \wedge B \vdash C$. By the deduction theorem it follows that $A \to (B \to C) \vdash A \wedge B \to C$. And, again by the deduction theorem, it also follows that $\vdash (A \to (B \to C)) \to (A \wedge B \to C)$. The reader would find it a difficult exercise to construct in a direct way (i.e., without applying the deduction theorem or using its method of proof) a logical proof of $(A \to (B \to C)) \to (A \wedge B \to C)$.

In general, it is much easier to show that $A_1, \dots, A_{n-1},\ A \vdash B$ than to show that $A_1, \dots, A_{n-1} \vdash A \to B$. The deduction theorem is a simple way to show the existence of certain (logical) deductions without having to exhibit those logical deductions explicitly. It is easy to write down a logical deduction of C from $A \to (B \to C)$ and $A \wedge B$; so, $A \to (B \to C),\ A \wedge B \vdash C$. Then, by two applications of the deduction theorem, one knows that $\vdash (A \to (B \to C)) \to (A \wedge B \to C)$, without having to write down a logical proof of the latter formula, which would be a rather complicated job. Following the proof of the deduction theorem one is able in principle to exhibit such a logical proof, but in most cases we are not interested in writing down this (logical) proof explicitly.

It is possible to derive additional results which make it easy to show that certain deductions exist without having to write down those deductions explicitly. One result is called *Reductio ad absurdum*; it says that in order to deduce $\neg A$ (from Γ, where Γ is a finite list of zero or more formulas) it suffices to deduce a contradiction (B and $\neg B$) from the assumption A (together with Γ). Another result is called $\vee$-*elimination*: in order to deduce C from $A \vee B$ (and Γ), it suffices to deduce C from A (and Γ) and to deduce C from B (and Γ).

The proof system of Section 2.6 contains only one rule, Modus Ponens. However, many other rules can be derived, for example, the rule called $\wedge$-*introduction*: from the two formulas A and B one can deduce the one formula $A \wedge B$. This result is obtained by using the axiom $A \to (B \to A \wedge B)$ and two applications of Modus Ponens. The next theorem contains the results just mentioned and a number of related similar results.

Theorem 2.25 (Introduction and Elimination Rules). *For any finite list Γ of (zero or more) formulas, and for any formulas A, B, C:*

	INTRODUCTION	*ELIMINATION*
$\to$	*If* $\Gamma, A \vdash B$, *then* $\Gamma \vdash A \to B$	$A, A \to B \vdash B$
$\wedge$	$A, B \vdash A \wedge B$	$A \wedge B \vdash A$ $A \wedge B \vdash B$
$\vee$	$A \vdash A \vee B$ $B \vdash A \vee B$	*If* $\Gamma, A \vdash C$ *and* $\Gamma, B \vdash C$, *then* $\Gamma, A \vee B \vdash C$
$\neg$	*If* $\Gamma, A \vdash B$ *and* $\Gamma, A \vdash \neg B$, *then* $\Gamma \vdash \neg A$ *(reductio ad absurdum)*	$\neg\neg A \vdash A$ *(double negation elimination)* $A, \neg A \vdash B$ *(weak negation elimination)*
$\rightleftarrows$	$A \to B, B \to A \vdash A \rightleftarrows B$	$A \rightleftarrows B \vdash A \to B$ $A \rightleftarrows B \vdash B \to A$

Proof. $\to$*-introduction* is the deduction theorem.

$\to$*-elimination*, $\wedge$*-introduction*, $\wedge$*-elimination*, $\vee$*-introduction*, *double negation elimination* and the three $\rightleftarrows$*-rules* are done in Exercise 2.39.

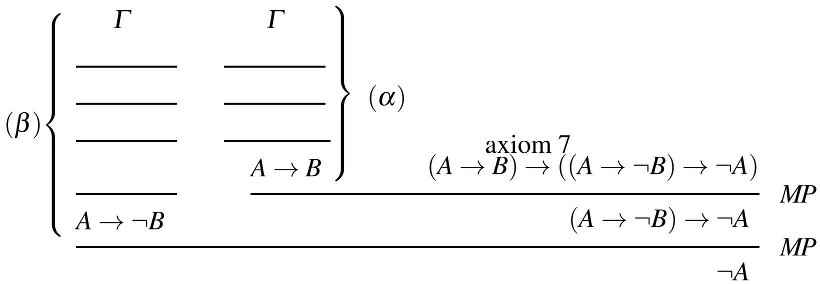
$\vee$*-elimination*: Suppose $\Gamma, A \vdash C$ and $\Gamma, B \vdash C$. Then by the deduction theorem $\Gamma \vdash A \to C$ and $\Gamma \vdash B \to C$. The following schema shows that $A \to C$, $B \to C$, $A \vee B \vdash C$:

$$\dfrac{B \to C \qquad \dfrac{A \to C \qquad \overset{\text{axiom 6}}{(A \to C) \to ((B \to C) \to (A \vee B \to C))}}{(B \to C) \to (A \vee B \to C)}\,MP}{\dfrac{A \vee B \to C \qquad A \vee B}{C}\,MP}\,MP$$

Hence, Γ, $A \vee B \vdash C$.

Weak negation elimination: Evidently, (1) A, $\neg A$, $\neg B \vdash A$, and (2) A, $\neg A$, $\neg B \vdash \neg A$. From (1) and (2) it follows by $\neg$-introduction that (3) A, $\neg A \vdash \neg\neg B$. And, by double negation elimination, also (4) $\neg\neg B \vdash B$. From (3) and (4) it follows that A, $\neg A \vdash B$. By this rule, from a contradiction A, $\neg A$, any formula B can be deduced.

$\neg$*-introduction* (reductio ad absurdum): Suppose $\Gamma, A \vdash B$ and $\Gamma, A \vdash \neg B$. Then by the deduction theorem $\Gamma \vdash A \to B$ and $\Gamma \vdash A \to \neg B$. Let (α) be a deduction of $A \to B$ from Γ and let (β) be a deduction of $A \to \neg B$ from Γ. Then the schema below is a deduction of $\neg A$ from Γ.



□

Exercise 2.45. Show that $A \wedge B \to C \vdash A \to (B \to C)$.

Exercise 2.46. Show that $\vdash (A \to B) \to (A \to ((B \to C) \to C))$.

Exercise 2.47. Show: if $A_1,\ A_2 \vdash B$, then $\vdash A_1 \wedge A_2 \to B$.

Exercise 2.48. Show that: If $\vdash (A_1 \wedge A_2) \wedge A_3 \to B$, then $A_1, A_2, A_3 \vdash B$.

Exercise 2.49. Prove or refute without making use of the completeness theorem: If $\vdash A \to C$ and $\vdash B \to C$, then $A \vee B \vdash C$. You may make use of the logical axiom $(A \to C) \to ((B \to C) \to (A \vee B \to C))$.

Exercise 2.50. Using $\vee$-elimination, show that $A \vee B,\ B \to C \vdash A \vee C$.

Exercise 2.51. Use $\neg$-introduction to show: if $A \vdash B$, then $\neg B \vdash \neg A$.

Exercise 2.52. Using $\neg$-introduction and exercise 2.51, show that $\vdash A \vee \neg A$.

Exercise 2.53. Using $\vee$-elimination, $\neg$-introduction and weak negation elimination, show that $\neg A,\ \neg B \vdash \neg(A \vee B)$.

Exercise 2.54. Use $\neg$-introduction to show: if $A \vdash \neg A$, then $\vdash \neg A$.

Exercise 2.55. Prove or refute (by means of a counterexample): for all formulas A, B, if $\vdash A \vee B$, then $\vdash A$ or $\vdash B$. Carefully specify your arguments.

Exercise 2.56. Prove or refute (by means of a counterexample): for all formulas A, if not $\vdash A$, then $\vdash \neg A$. Carefully specify your arguments and do not use the completeness theorem.

Exercise 2.57. Prove or refute, carefully specifying your arguments and not making use of the completeness theorem:
a) If $\vdash A \to B$, then $A \vdash B$. b) If $\vdash \neg A$, then not $\vdash A$.

Exercise 2.58. Show that $\neg A \vee B \vdash A \to B$. Next show: a) $A \to B,\ \neg(\neg A \vee B) \vdash \neg\neg A$ and b) $A \to B,\ \neg(\neg A \vee B) \vdash \neg A$. Conclude from a) and b) by $\neg$-introduction that $A \to B \vdash \neg\neg(\neg A \vee B)$ and hence $A \to B \vdash \neg A \vee B$.

Exercise 2.59 (Completeness). In this exercise we shall prove the *completeness theorem* for classical propositional logic along the lines of L. Kalmár, 1934-5.

Consider the truth table for a formula $E(A,B)$ built from the formulas A and B. To each entry (or line) of this truth table a corresponding deducibility relationship holds, as indicated below:

	A	B	$E(A,B)$	
u_1	1	1	$u_1(E)$	$A,\ B \vdash E_1^*$
u_2	1	0	$u_2(E)$	$A, \neg B \vdash E_2^*$
u_3	0	1	$u_3(E)$	$\neg A,\ B \vdash E_3^*$
u_4	0	0	$u_4(E)$	$\neg A, \neg B \vdash E_4^*$

where $E_i^* = E$ if $u_i(E) = 1$ and $E_i^* = \neg E$ if $u_i(E) = 0$ ($i = 1, 2, 3, 4$).

a) Establish the first two deducibility relationships for $E = A \wedge B$ and the last two for $E = A \vee B$.

b) Using the result mentioned above prove the completeness theorem for classical propositional logic: if $\models E$, then $\vdash E$.

2.7.2 *Natural Deduction**

Hilbert's proof system, presented in Section 2.6, has several axiom schemas and only one rule, Modus Ponens. In his *Untersuchungen über das logische Schliessen* G. Gentzen [9] introduced a different, but equivalent, proof system which has several rules, but no axioms. This proof system is called Gentzen's system of *Natural Deduction*. Logical proofs in this system are very similar to the informal proofs in daily reasoning, which makes the search for a logical proof in this system much easier than in a Hilbert-type proof system. Before the rules are presented some of them will be discussed and the notation explained.

→-Introduction: Suppose B is derived from the assumption A (and perhaps other assumptions as well); notation:

$$\begin{array}{c} A \\ \vdots \\ B \end{array}$$

Then one can derive $A \to B$, cancelling the assumption A; notation:

$$\dfrac{\begin{array}{c} [A]^i \\ \vdots \\ B \end{array}}{A \to B}\, i \qquad \text{where } i \text{ is a natural number.}$$

Note that this rule corresponds to the deduction theorem (Theorem 2.24).

¬-Introduction: Suppose a contradiction (B and $\neg B$) is derived from one or more assumption formulas among which is A. Notation:

$$\begin{array}{c} A \\ \vdots \\ B \quad \neg B \end{array}$$

Then one can obtain a deduction of $\neg A$ from the assumptions without A. Notation:

$$\dfrac{\begin{array}{c} [A]^i \\ \vdots \\ B \quad \neg B \end{array}}{\neg A}\, i$$

$\vee$-*Elimination* : Suppose one has a deduction of C from the assumption A and another deduction of the same formula C from the assumption B, where in both cases other assumptions may be present. Then one can obtain a deduction of C from the assumption $A \vee B$, cancelling the assumptions A and B. Notation:

$$\dfrac{A \vee B \qquad \begin{array}{c} [A]^i \\ \vdots \\ C \end{array} \qquad \begin{array}{c} [B]^i \\ \vdots \\ C \end{array}}{C}\, i$$

Having explained how to read the more complicated rules of natural deduction, below all Gentzen rules for *natural deduction* are presented.

GENTZEN'S INTRODUCTION RULES		GENTZEN'S ELIMINATION RULES	
$\&I$	$\dfrac{A \quad B}{A \wedge B}$	$\&E$	$\dfrac{A \wedge B}{A} \qquad \dfrac{A \wedge B}{B}$
$\vee I$	$\dfrac{A}{A \vee B} \qquad \dfrac{B}{A \vee B}$	$\vee E$	$\dfrac{A \vee B \qquad \begin{array}{c} [A] \\ \vdots \\ C \end{array} \qquad \begin{array}{c} [B] \\ \vdots \\ C \end{array}}{C}$
$\to I$	$\dfrac{\begin{array}{c} [A] \\ \vdots \\ B \end{array}}{A \to B}$	$\to E$	$\dfrac{A \quad A \to B}{B}$
$\neg I$	$\dfrac{\begin{array}{c} [A] \\ \vdots \\ B \quad \neg B \end{array}}{\neg A}$	$w\,\neg E$ / $d\,\neg E$	$\dfrac{A \quad \neg A}{B}$ (w = weak) $\qquad \dfrac{\neg\neg A}{A}$ (d = double)

The reader should note the analogy with the Introduction and Elimination rules in Theorem 2.25, but he should also see the difference. For instance, $A \vdash A \vee B$ says that $A \vee B$ can be obtained from A and the logical axioms by applying the rule Modus Ponens a finite number of times, while $\dfrac{A}{A \vee B}$ itself is a rule of inference in the natural deduction system, as Modus Ponens is a rule of inference in the axiomatic system of Section 2.6. In other words, $A \vdash A \vee B$ says that $\dfrac{A}{A \vee B}$ is a *derived rule* of inference in the axiomatic system of Section 2.6.

Example 2.18. Below are some examples of deductions in Gentzen's system of Natural Deduction.

(i) $(A \to B) \to ((B \to C) \to (A \to C))$

$$
(3)\;\dfrac{(2)\;\dfrac{(1)\;\dfrac{\dfrac{\dfrac{{}^{1}[A] \qquad [A \to B]^{3}}{B}\,\to E \qquad [B \to C]^{2}}{C}\,\to E}{A \to C}\,\to I}{(B \to C) \to (A \to C)}\,\to I}{(A \to B) \to ((B \to C) \to (A \to C))}\,\to I
$$

The reader should note the analogy with the way in which we intuitively verify that $(A \to B) \to ((B \to C) \to (A \to C))$ is true.
To show: $(A \to B) \to ((B \to C) \to (A \to C))$.
So suppose $A \to B$; then to show $(B \to C) \to (A \to C)$.
So suppose $B \to C$; then to show $A \to C$.
So suppose A; then to show C.
Now from A and $A \to B$ it follows that B. And from B and $B \to C$ it follows that C. So C follows from A, $B \to C$ and $A \to B$. Hence $A \to C$ follows from $B \to C$ and $A \to B$. Therefore $(B \to C) \to (A \to C)$ follows from $A \to B$. Consequently, $(A \to B) \to ((B \to C) \to (A \to C))$.

(ii) $\neg\neg A \to A$

$$
(1)\;\dfrac{\dfrac{[\neg\neg A]^{1}}{A}\,d\neg E}{\neg\neg A \to A}\,\to I
$$

(iii) $A \to \neg\neg A$

$$
(2)\;\dfrac{(1)\;\dfrac{{}^{2}[A] \qquad [\neg A]^{1}}{\neg\neg A}\,\neg I}{A \to \neg\neg A}\,\to I
$$

(iv) In the deduction of $A \vee \neg A$ below, the reader should again note the analogy with the way in which we intuitively show that $A \vee \neg A$ is true. Suppose that $\neg(A \vee \neg A)$. Then, since $A \vee \neg A$ follows from A, $\neg A$. But also, since $A \vee \neg A$ follows from $\neg A$, $\neg(\neg A)$. So from $\neg(A \vee \neg A)$ it follows that both $\neg A$ and $\neg(\neg A)$. Therefore, by $\neg$-introduction, $\neg\neg(A \vee \neg A)$ and hence, by double $\neg$-elimination, $A \vee \neg A$.

$$\dfrac{(3)\ \dfrac{(1)\ \dfrac{{}^3[\neg(A \vee \neg A)] \qquad \dfrac{[A]^1}{A \vee \neg A}\ \vee I}{\neg A}\ \neg I \qquad (2)\ \dfrac{{}^3[\neg(A \vee \neg A)] \qquad \dfrac{[\neg A]^2}{A \vee \neg A}\ \vee I}{\neg\neg A}\ \neg I}{\neg\neg(A \vee \neg A)}\ \neg I}{A \vee \neg A}\ d\neg E$$

Definition 2.12 (Deducibility in natural deduction). a) Let Γ be a (possibly infinite) set of formulas. B is *deducible from* Γ in Gentzen's system of Natural Deduction := B can be obtained by one or more (but finitely many) applications of Gentzen's rules of natural deduction from uncancelled assumptions that belong to the set Γ. **Notation**: $\Gamma \vdash_{ND} B$.
b) In case Γ is empty, we say that B is *provable* in Gentzen's system of natural deduction. **Notation**: $\vdash_{ND} B$.

Example 2.19. In Example 2.18 we have seen:

$A \to B \vdash_{ND} (B \to C) \to (A \to C)$	$\vdash_{ND} (A \to B) \to ((B \to C) \to (A \to C))$
$\neg\neg A \vdash_{ND} A$	$\vdash_{ND} \neg\neg A \to A$
$A \vdash_{ND} \neg\neg A$	$\vdash_{ND} A \to \neg\neg A$
	$\vdash_{ND} A \vee \neg A$

Once having shown Theorem 2.25 (introduction and elimination rules), one easily sees that Gentzen's system of natural deduction is equivalent to the axiomatic (Hilbert-type) system of Section 2.6.

Theorem 2.26. *$\Gamma \vdash B$ iff $\Gamma \vdash_{ND} B$.*

Proof. i) Suppose $\Gamma \vdash B$. One easily checks that all the axioms of (classical) propositional logic are provable in Gentzen's system of natural deduction. Modus Ponens *MP* is precisely Gentzen's rule $\to E$. It follows that $\Gamma \vdash_{ND} B$.
ii) Suppose $\Gamma \vdash_{ND} B$. a) If B is an element of Γ, then $\Gamma \vdash B$.
b) Theorem 2.25 shows that all steps made in Gentzen's rules of natural deduction are also available for the notion of (Hilbert-type) deducibility of Section 2.6. More precisely, Gentzen's rule $\vee E$, for instance, says that if $\Delta, A \vdash_{ND} C$ and $\Delta, B \vdash_{ND} C$, then $\Delta, A \vee B \vdash_{ND} C$ for any set Δ of formulas. Now suppose (by induction hypothesis) that $\Delta, A \vdash C$ and $\Delta, B \vdash C$; then by $\vee$-elimination in Theorem 2.25, $\Delta, A \vee B \vdash C$. By a) and b) it follows (by induction on the length of a given ND-deduction of B from Γ in Gentzen's system of natural deduction) that $\Gamma \vdash B$. □

Exercise 2.60. Show that: i) $\neg(A \wedge B) \vdash_{ND} \neg A \vee \neg B$, ii) $\neg A \vee \neg B \vdash_{ND} \neg(A \wedge B)$. Keep in mind the way in which we would intuitively verify that the conclusion follows from the premisses.

Exercise 2.61. i) Show that $A \vdash B \to A$ and follow the proof of Theorem 2.26, part i), to convert the given deduction of $B \to A$ from A in Hilbert's system into a deduction of $B \to A$ from A in Gentzen's system of natural deduction.
ii) Show that $A \to B \vdash_{ND} \neg B \to \neg A$ and follow the proof of Theorem 2.26, part ii) to show that $A \to B \vdash \neg B \to \neg A$.

2.8 Tableaux

In this section we will introduce another notion of provability and of deducibility, which is based on the work of E. Beth [2] and of G. Gentzen [9], and equivalent to the corresponding notions defined in Section 2.6. The advantage of Beth's and Gentzen's notions is that the search for a deduction of B from $A_1, \ldots, A_n$ becomes a mechanical matter and is not achieved by the method of trial and error, as is (sometimes) the case for the historically older notions of Section 2.6, which are essentially based on the work of G. Frege [7] (1848-1925) and B. Russell [25] (1872-1970). This advantage is obtained by reducing the number of axiom-schemes to one, essentially $A \to A$, and by replacing the axioms by T and F rules, two for each connective. The presentation chosen here is close to the one of R. Smullyan [23] and was introduced by M. Fitting [6].

Definition 2.13 (Signed formula). A *signed formula* is any expression of the form $T(A)$ or $F(A)$, where A is a formula.

In the case of classical logic, the intended meanings of $T(A)$ and $F(A)$, in Beth's semantic tableaux rules, are as follows: $T(A)$: A is true, $F(A)$: A is false. (The intended meanings of $T(A)$ and $F(A)$ for modal and intuitionistic logic are different.)

If it is clear from the context what is meant, we will simply write TA instead of $T(A)$ and FA instead of $F(A)$. For instance, instead of $T(B \wedge C)$ we will mostly write $T\ B \wedge C$.

Definition 2.14 (Sequent). A *sequent* S is any finite set of signed formulas.

For example, $\{T\ P_1 \to P_2,\ F\ \neg P_1 \wedge P_2,\ F\ \neg P_2 \vee (P_2 \to P_1)\}$ is a sequent. In Gentzen's approach the intended meaning of a sequent $\{TB_1, \ldots, TB_m,\ FC_1, \ldots, FC_n\}$ is as follows: if B_1 and ... and B_m, then C_1 or ... or C_n.

Below we present the T- and F- *tableaux rules for classical propositional logic*; next we will explain how to read them, either as *semantic tableaux rules* in the sense of Beth or as *Gentzen-type rules*. In what follows, S will always denote a sequent.

$T\wedge$	$S,\ T\ B\wedge C$ $S,\ TB,\ TC$	$F\wedge$	$S,\ F\ B\wedge C$ $S,\ FB\ \mid\ S,\ FC$
$T\vee$	$S,\ T\ B\vee C$ $S,\ TB\ \mid\ S,\ TC$	$F\vee$	$S,\ F\ B\vee C$ $S,\ FB,\ FC$
$T\rightarrow$	$S,\ T\ B\rightarrow C$ $S,\ FB\ \mid\ S,\ TC$	$F\rightarrow$	$S,\ F\ B\rightarrow C$ $S,\ TB,\ FC$
$T\neg$	$S,\ T\ \neg B$ $S,\ FB$	$F\neg$	$S,\ F\ \neg B$ $S,\ TB$

Notation: $S,\ TA$ stands for $S\cup\{TA\}$, i.e., the set containing all signed formulas in S and in addition TA; and $S,\ FA$ similarly stands for $S\cup\{FA\}$. Instead of $\{TB_1,\ldots,TB_m,\ FC_1,\ldots,FC_n\}$ we often simply write $TB_1,\ldots,TB_m,\ FC_1,\ldots,FC_n$. For example, by $\{TD,\ FE\},\ TA$ we mean $\{TD,\ FE,\ TA\}$, but we will usually write $TD,\ FE,\ TA$.

Since $S,\ T\ B\wedge C$ stands for $S\cup\{T\ B\wedge C\}$, and since this latter set is equal to $S\cup\{T\ B\wedge C,\ T\ B\wedge C\}$, the following rule

$$\begin{array}{c} S,\ T\ B\wedge C \\ S,\ T\ B\wedge C,\ TB,\ TC \end{array}$$

is a derived rule. So, in any application of any rule the T-signed or the F-signed formula to which the rule is applied may be repeated in the lower half of the rule.

Beth's semantic tableaux rules The rules given above can be read in two ways.

First, *read downwards*, as *semantic tableaux rules* in the sense of E. Beth, *interpreting the signed formulas* rather than the sequents. For example, in the case of rule $T\rightarrow$: if $B\rightarrow C$ is true ($T\ B\rightarrow C$), then there are two possibilities, B is false (FB) *or* C is true (TC). And in the case of rule $F\rightarrow$: if $B\rightarrow C$ is false ($F\ B\rightarrow C$), then B is true (TB) and C is false (FC).

This way of reading the rules is derived from E. Beth's [2] method of semantic tableaux. A formula B is called *tableau-deducible from* given formulas $A_1,\ldots,A_n$ if it turns out to be impossible that $A_1,\ldots,A_n$ are all 1 and B is 0; more precisely, if all sequents which result from application of the rules to the supposition $TA_1,\ldots,TA_n$, FB ($A_1,\ldots,A_n$ are all 1 and B is 0) and to which no further rules can be applied, turn out to be contradictory, i.e., for all such sequents there is an atomic formula P such that both TP (P is true) and FP (P is false) occur in it (see Def. 2.16 and 2.18).

Note that we essentially have used this idea in exercise 2.11 to verify that, for instance, $\models (P\rightarrow Q)\rightarrow(\neg Q\rightarrow\neg P)$ or, equivalently, $(P\rightarrow Q)\models(\neg Q\rightarrow\neg P)$, by showing that it is impossible that in some line of the truth table $(P\rightarrow Q)$ is 1 and $(\neg Q\rightarrow\neg P)$ is 0. In the left column of Example 2.20 we apply the tableaux rules to $T\ (P\rightarrow Q),\ F\ (\neg Q\rightarrow\neg P)$ and in the right column of Example 2.20 we give the interpretation of the left column in the sense of E. Beth.

Example 2.20.

	Suppose in some line of its truth table
$T\ (P \to Q),\ F(\neg Q \to \neg P)$	$(P \to Q)$ is 1 and $\neg Q \to \neg P$ is 0.
$T\ P \to Q,\ T\neg Q,\ F\neg P$	Then $P \to Q$ is 1, $\neg Q$ is 1 and $\neg P$ is 0.
$T\ P \to Q,\ FQ,\ F\neg P$	So, $P \to Q$ is 1, Q is 0 and $\neg P$ is 0 in that line.
$T\ P \to Q,\ FQ,\ TP$	So, $P \to Q$ is 1, Q is 0 and P is 1 in that line.
$FP,\ FQ,\ TP \mid TQ,\ FQ,\ TP$	So, P is 0, Q is 0 and P is 1, or,
	Q is 1, Q is 0 and P is 1 in that same line.
	And both are impossible.

Informally, we say that the left column in Example 2.20 is a *tableau* $\mathscr{T}$ with initial branch $\mathscr{B}_0 = \{T\ (P \to Q),\ F(\neg Q \to \neg P)\}$. This tableau $\mathscr{T}$ consists of two tableau branches $\mathscr{B}_{31}$ and $\mathscr{B}_{32}$, with $\mathscr{B}_{31} = \{T\ (P \to Q),\ F(\neg Q \to \neg P),\ T\neg Q,\ F\neg P,\ FQ,\ TP,\ FP\}$, containing all signed formulas in the left half of the tableau and $\mathscr{B}_{32} = \{T\ (P \to Q),\ F(\neg Q \to \neg P),\ T\neg Q,\ F\neg P,\ FQ,\ TP,\ TQ\}$, containing all signed formulas in the right half of the tableau. The branch $\mathscr{B}_{31}$ is *closed* because it contains TP and FP, and the branch $\mathscr{B}_{32}$ is *closed* because it contains TQ and FQ. Both branches are *completed*, i.e., for each signed formula in the branch the corresponding T- or F-rule has been applied.

Definition 2.15 ((Tableau) Branch). (a) A *tableau branch* is a set of signed formulas. A branch is *closed* if it contains signed formulas TA and FA for some formula A. A branch that is not closed is called *open*.
(b) Let $\mathscr{B}$ be a branch and TA, resp. FA, a signed formula occurring in $\mathscr{B}$. TA, resp. FA, is *fulfilled* in $\mathscr{B}$ if (i) A is atomic, or (ii) $\mathscr{B}$ contains the bottom formulas in the application of the corresponding rule to A, and in case of the rules $T\vee$, $F\wedge$ and $T \to$, $\mathscr{B}$ contains one of the bottom formulas in the application of these rules.
(c) A branch $\mathscr{B}$ is *completed* if $\mathscr{B}$ is closed or every signed formula in $\mathscr{B}$ is fulfilled in $\mathscr{B}$.

More formally, in Example 2.20 we call $\mathscr{B}_0 = \{T\ (P \to Q),\ F(\neg Q \to \neg P)\}$ the initial branch and $\mathscr{T}_0 = \{\mathscr{B}_0\}$ a tableau (with initial branch $\mathscr{B}_0$).

Let $\mathscr{B}_1 = \{T\ (P \to Q),\ F(\neg Q \to \neg P),\ T\neg Q,\ F\neg P\}$. Then $\mathscr{T}_1 = \{\mathscr{B}_1\}$ is called a *one-step expansion* of $\mathscr{T}_0$, because there is a signed formula in $\mathscr{B}_0$, to wit $F(\neg Q \to \neg P)$, such that $\mathscr{B}_1 = \mathscr{B}_0 \cup \{T\neg Q,\ F\neg P\}$.

Let $\mathscr{B}_2 = \{T\ (P \to Q),\ F(\neg Q \to \neg P),\ T\neg Q,\ F\neg P,\ FQ\}$. Then $\mathscr{T}_2 = \{\mathscr{B}_2\}$ is again a *one-step expansion* of $\mathscr{T}_1$.

Let $\mathscr{B}_3 = \{T\ (P \to Q),\ F(\neg Q \to \neg P),\ T\neg Q,\ F\neg P,\ FQ,\ TP\}$. Then $\mathscr{T}_3 = \{\mathscr{B}_3\}$ is a *one-step expansion* of $\mathscr{T}_2$.

Finally, let $\mathscr{B}_{31} = \{T\ (P \to Q),\ F(\neg Q \to \neg P),\ T\neg Q,\ F\neg P,\ FQ,\ TP,\ FP\}$ and $\mathscr{B}_{32} = \{T\ (P \to Q),\ F(\neg Q \to \neg P),\ T\neg Q,\ F\neg P,\ FQ,\ TP,\ TQ\}$. Then $\mathscr{T}_4 = \{\mathscr{B}_{31},\ \mathscr{B}_{32}\}$ is called a *one-step expansion* of $\mathscr{T}_3$, because there is a signed formula in $\mathscr{B}_3$, to wit $T\ (P \to Q)$, such that $\mathscr{B}_{31} = \mathscr{B}_3 \cup \{FP\}$ and $\mathscr{B}_{32} = \mathscr{B}_3 \cup \{TQ\}$.

$\mathscr{T}_0$, $\mathscr{T}_1$, $\mathscr{T}_2$, $\mathscr{T}_3$ and $\mathscr{T}_4$ are all tableaux with initial branch $\mathscr{B}_0$.

The branches $\mathscr{B}_0$, $\mathscr{B}_1$, $\mathscr{B}_2$ and $\mathscr{B}_3$ are not closed and not completed. But the branches $\mathscr{B}_{31}$ and $\mathscr{B}_{32}$ are completed and both are also closed.

We shall call, for instance, $\mathscr{T}_3 = \{\mathscr{B}_3\}$ a *tableau* with initial branch or sequent $\mathscr{B}_0$, because there is a sequence $\mathscr{T}_0, \mathscr{T}_1, \ldots, \mathscr{T}_3$ such that $\mathscr{T}_0 = \{\mathscr{B}_0\}$ and each $\mathscr{T}_{i+1}$ is a one-step expansion of $\mathscr{T}_i$ $(0 \leq i < 3)$. This tableau $\mathscr{T}_3$ (with initial branch $\mathscr{B}_0$) is not yet completed, because its only branch $\mathscr{B}_3$ is not completed: the $T \rightarrow$ rule has not yet been applied to $T(P \rightarrow Q)$. And $\mathscr{T}_3 = \{\mathscr{B}_3\}$ is open, because it contains an open branch, to wit $\mathscr{B}_3$ itself. The tableau $\mathscr{T}_4 = \{\mathscr{B}_{31}, \mathscr{B}_{32}\}$, however, is completed, because each of its branches is completed and also closed, because all its branches are closed.

Definition 2.16 (Tableau). (a) A set of branches $\mathscr{T}$ is a *tableau* with initial branch $\mathscr{B}_0$ if there is a sequence $\mathscr{T}_0, \mathscr{T}_1, \ldots, \mathscr{T}_n$ such that $\mathscr{T}_0 = \{\mathscr{B}_0\}$, each $\mathscr{T}_{i+1}$ is a one-step expansion of $\mathscr{T}_i$ $(0 \leq i < n)$ and $\mathscr{T} = \mathscr{T}_n$.
(b) We say that a finite $\mathscr{B}$ has tableau $\mathscr{T}$ if $\mathscr{T}$ is a tableau with initial branch $\mathscr{B}$.
(c) A tableau $\mathscr{T}$ is *open* if some branch $\mathscr{B}$ in it is open, otherwise $\mathscr{T}$ is *closed*.
(d) A tableau is *completed* if each of its branches is completed, i.e., no application of a tableau rule can change the tableau.

Example 2.21.
We make a tableau starting with $T(P \rightarrow Q),\ F(P \wedge Q)$:

$$\begin{array}{c} T(P \rightarrow Q),\ F(P \wedge Q) \\ FP,\ F(P \wedge Q) \mid TQ,\ F(P \wedge Q) \\ FP,\ FP \mid FP,\ FQ \mid TQ,\ FP \mid TQ,\ FQ \end{array}$$

Let $\mathscr{B}_1$ be the leftmost branch, consisting of the formulas $T(P \rightarrow Q),\ F(P \wedge Q),\ FP$ and FP, i.e., $\mathscr{B}_1 = \{T(P \rightarrow Q),\ F(P \wedge Q),\ FP,\ FP\}$. Let $\mathscr{B}_2$ be the second branch from the left, so $\mathscr{B}_2 = \{T(P \rightarrow Q),\ F(P \wedge Q),\ FP,\ FQ\}$. Let $\mathscr{B}_3$ be the third branch from the left, so $\mathscr{B}_3 = \{T(P \rightarrow Q),\ F(P \wedge Q),\ TQ,\ FP\}$. Finally, let $\mathscr{B}_4$ be the rightmost branch, i.e., $\mathscr{B}_4 = \{T(P \rightarrow Q),\ F(P \wedge Q),\ TQ,\ FQ\}$.

Then $\mathscr{T} = \{\mathscr{B}_1, \mathscr{B}_2, \mathscr{B}_3, \mathscr{B}_4\}$ is a tableau with $\mathscr{B}_0 = \{T(P \rightarrow Q), F(P \wedge Q)\}$ as initial branch. Branch $\mathscr{B}_4$ is completed and closed, because it contains TQ and FQ. The branches $\mathscr{B}_1, \mathscr{B}_2, \mathscr{B}_3$ are completed and open. Hence, the tableau $\mathscr{T} = \{\mathscr{B}_1, \mathscr{B}_2, \mathscr{B}_3, \mathscr{B}_4\}$ is completed, because all of its branches are completed and the tableau $\mathscr{T}$ is open, since at least one of its branches is open.

From the formulation of the tableaux rules, we see immediately that our tableaux have the so-called *subformula property*: each formula in any sequent of a tableau is a subformula of some formula occurring in the preceding sequents. For that reason, any tableau (in classical propositional logic) is necessarily a *finite* sequence of sequents. For instance, all formulas in the tableau in Example 2.20 are subformulas of $P \rightarrow Q$ and/or $\neg Q \rightarrow \neg P$.

From the examples in Section 2.6 it is clear that a Hilbert-type proof system does not have the subformula property. For instance, we have given a deduction of $A \rightarrow C$ from $A \rightarrow B$ and $B \rightarrow C$; in this deduction we have used the formula $A \rightarrow (B \rightarrow C)$ and even more complex ones, which are subformulas of neither the premisses nor the conclusion. Modus Ponens is responsible for this: E may be deduced from D and $D \rightarrow E$; but $D \rightarrow E$ is not a subformula of E and D is not necessarily one.

Definition 2.17 (Tableau-deduction). **(a)** A (logical) *tableau-deduction* of B *from* $A_1,\dots,A_n$ (in propositional logic) is a tableau $\mathscr{T}$ with $\mathscr{B}_0 = \{TA_1,\dots,TA_n,FB\}$ as initial branch, such that all branches of $\mathscr{T}$ are closed.

In case $n = 0$, i.e., there are no premisses $A_1,\dots,A_n$, this definition reduces to:
(b) A (logical) *tableau-proof* of B (in classical propositional logic) is a tableau $\mathscr{T}$ with $\mathscr{B}_0 = \{FB\}$ as initial sequent, such that all branches of $\mathscr{T}$ are closed.

Example 2.22. **(a)** The following is a tableau-deduction of $\neg P \vee \neg Q$ from $\neg(P \wedge Q)$.

$$\begin{array}{c} T\ \neg(P\wedge Q),\ F\ \neg P\vee\neg Q \\ F\ P\wedge Q,\ F\ \neg P\vee\neg Q \\ F\ P\wedge Q,\ F\ \neg P,\ F\ \neg Q \\ F\ P\wedge Q,\ TP,\ F\ \neg Q \\ F\ P\wedge Q,\ TP,\ TQ \\ FP,\ TP,\ TQ \mid FQ,\ TP,\ TQ \end{array}$$

(b) The following is a tableau-proof of $((P \to Q) \to P) \to P$, i.e., Peirce's law.

$$\begin{array}{c} F\ ((P\to Q)\to P)\to P \\ T\ (P\to Q)\to P,\ FP \\ F\ P\to Q,\ FP \mid TP,\ FP \\ TP,\ FQ,\ FP \mid \end{array}$$

Definition 2.18 (Tableau-deducible). **(a)** B is *tableau-deducible from* $A_1,\dots,A_n$ (in classical propositional logic) if there exists a tableau-deduction of B from $A_1,\dots,A_n$.
Notation: $A_1,\dots,A_n \vdash' B$. By $A_1,\dots,A_n \nvdash' B$ we mean: not $A_1,\dots,A_n \vdash' B$.
(b) B is *tableau-provable* (in classical propositional logic) if there exists a tableau-proof of B. **Notation**: $\vdash' B$.
(c) For Γ a (possibly infinite) set of formulas, B is *tableau-deducible from* Γ if there exists a finite list $A_1,\dots,A_n$ of formulas in Γ such that $A_1,\dots,A_n \vdash' B$.
Notation: $\Gamma \vdash' B$.

Example 2.23. **(a)** $\neg(P \wedge Q)\ \vdash' \neg P \vee \neg Q$, because in Example 2.22 (a) we have given a tableau-deduction of $\neg P \vee \neg Q$ from $\neg(P \wedge Q)$. One also easily checks that, equivalently, $\vdash' \neg(P \wedge Q) \to \neg P \vee \neg Q$.
(b) $\vdash' ((P \to Q) \to P) \to P$, because in Example 2.22 (b) we have given a tableau-proof of $((P \to Q) \to P) \to P$. One also easily checks that, equivalently, $(P \to Q) \to P \vdash' P$.

Note that by our definitions $A \vdash' B$ is trivially equivalent to $\vdash' A \to B$ (because a tableau starting with $F\ A \to B$ continues with TA, FB), while the corresponding result for $\vdash$ (Theorem 2.23 and 2.24) was not trivial at all.

It is important to note that the T- and F-rules and hence the notions of 'tableau-provable' and 'tableau-deducible from' are purely *syntactic*, i.e., they only refer to the forms of the formulas: for instance, rule $T\wedge$ tells us that any time we see an expression of the form $T\ B \wedge C$ we must write down the expressions TB and TC immediately below it; and a formula B is tableau-provable if starting with FB we end up with sequents which all contain both TP and FP for some atomic formula P.

Whether a formula B is tableau-provable or not only depends on the *form* of B, and precisely this justifies our use of the expression 'B is tableau-provable'.

So, we had good semantic reasons to choose the rules and the notions of 'tableau-provable' and 'tableau-deducible from' as they are, but once having these rules and these notions, we can forget the intuitive (semantic) motivation behind them and like a computer or machine/robot play with them in a purely syntactic way, i.e., apply the rules of the game, forgetting about their underlying ideas.

Gentzen-type rules A second way to read the T- and F- tableaux rules is to read them *upwards*, as *Gentzen-type rules*, interpreting the sequents rather than the signed formulas. Remember that a sequent $\{TA_1,\dots,TA_n,FB_1,\dots,FB_k\}$ is read as: if A_1 and ... and A_n, then B_1 or ... or B_k.

For example, taking $S = \{TD,\ FE\}$, rule $T\to$ becomes

$$\begin{array}{c} TD,\ FE,\ T\ B\to C \\ TD,\ FE,\ FB\ \mid\ TD,\ FE,\ TC \end{array}$$

and is read upwards as follows:

if	(*) D implies E or B	$(TD,\ FE,\ FB)$,
and	(**) D and C imply E	$(TD,\ FE,\ TC)$,
then	D and $B\to C$ imply E	$(TD,\ FE,\ T\ B\to C)$.

That rule $T\to$, read in this way, is intuitively correct is easily seen as follows: suppose (*), (**), D and $B\to C$; then by (*), E or B; if B, then by $B\to C$ also C; and hence by (**) E.

And again taking $S = \{TD,\ FE\}$, rule $F\to$ becomes

$$\begin{array}{c} TD,\ FE,\ F\ B\to C \\ TD,\ FE,\ TB,\ FC \end{array}$$

and is read upwards as follows:

if	(*) D and B imply E or C	$(TD,\ FE,\ TB,\ FC)$,
then	D implies E or $B\to C$	$(TD,\ FE,\ F\ B\to C)$.

That rule $F\to$, read in this way, is intuitively correct is seen as follows: suppose (*) and D; if $\neg B$, then $B\to C$ and hence E or $B\to C$; and if B, then D and B, and hence by (*), E or C; so, also E or $B\to C$.

This way of reading the rules is derived from G. Gentzen's system in [9]. Gentzen thought his rules reflected (the elementary steps in) the actual reasoning of human beings. With this reading the notion of tableau-provability is explained (see Def. 2.18) in terms of reducing a formula according to the rules to axioms essentially of the type $P\to P$. More precisely, a formula B is tableau-provable if $\{FB\}$ (to be read as $\to B$ or B) can be obtained by applying the rules to sequents of the form $\{\dots,TP,\ FP,\dots\}$ (to be read as: if ... and P, then P or ...), which can be conceived of as axioms.

Decidability Evidently, it is easy to decide whether a given sequence of symbols is a formula (of propositional logic). It is also easy to decide whether a given sequence of formulas is a (Hilbert-type) deduction (see Section 2.6) of a given formula B from given premisses $A_1,\dots,A_n$. And similarly, it is easy to decide whether a given tableau is a tableau-deduction of a given formula B from given premisses $A_1,\dots,A_n$.

But the question whether, given any formulas $A_1,\dots A_n$ and B, there *exists* a Hilbert-type deduction of B from $A_1,\dots A_n$, is not so easy to decide: one may search for such a deduction without finding one and this may be due to the fact that one is not smart enough – in which case one may continue trying to find one –, but also due to the fact that there is no such deduction – in which case one better stops searching. The deeper reason behind this is that Hilbert-type deductions do not have the subformula property: if one searches for a deduction of B from given premisses, one may try any formula D, not necessarily a subformula of the given formulas, in order to apply Modus Ponens to D and $D \to B$.

Interestingly, for any propositional formulas $A_1,\dots,A_n$, B, the question whether B is a valid (or logical) consequence of $A_1,\dots,A_n$ is *decidable*, i.e., there is a decision procedure (algorithm, mechanical test) which yields in finitely many steps an answer 'yes' or 'no': make the truth table of the formulas in question and check whether B is 1 in all lines where the premisses $A_1,\dots,A_n$ are all 1.

Similarly, for any propositional formulas $A_1,\dots,A_n$, B, the question whether there *exists* a tableau-deduction of B from given premisses $A_1,\dots,A_n$ is *decidable*, since there is a decision procedure which yields in finitely many steps an answer 'yes' or 'no': given $A_1,\dots,A_n$ and B, start a tableau with $\{TA_1,\dots,TA_n,FB\}$ as initial sequent and apply all possible tableau rules as frequently as possible; because of the subformula property, after finitely many steps the tableau will be finished; if all tableau branches are closed, then one has a tableau-deduction of B from $A_1,\dots,A_n$, and if some completed tableau branch is open, one can from any open completed tableau branch read off a line in the truth table in which $A_1,\dots,A_n$ are all 1 and B is 0, hence showing that $A_1,\dots,A_n \not\models B$. We shall prove this (completeness) result in Section 2.9, but will illustrate this result now with an example.

Example 2.24. We wonder whether from $P \to Q$ and $\neg P$ one may deduce $\neg Q$. So, we start a tableau with $\{TP \to Q,\ T\neg P,\ F\neg Q\}$:

$$\begin{array}{c} TP \to Q,\ T\neg P,\ F\neg Q \\ TP \to Q,\ FP,\ F\neg Q \\ TP \to Q,\ FP,\ TQ \\ FP, FP, TQ \mid TQ, FP, TQ \end{array}$$

For instance, the left tableau branch is completed but open, i.e. not closed. From it one may immediately read off a counterexample, i.e., a line in the truth table in which the premisses $P \to Q$ and $\neg P$ are 1 and $\neg Q$ is 0: corresponding with the occurrence of FP in the left completed tableau branch give P the value 0 and corresponding with the occurrence of TQ in the left completed tableau branch give Q the value 1.

P	Q	$P \to Q$	$\neg P$	$\neg Q$
0	1	1	1	0

This shows that $P \to Q, \neg P \not\models \neg Q$.

Once we have shown in Sect. 2.9 that the three notions $A_1,\ldots,A_n \models B$, $A_1,\ldots,A_n \vdash B$, and $A_1,\ldots,A_n \vdash' B$, although intensionally quite different, are equivalent, we have also a decision procedure for the question whether, given formulas $A_1,\ldots,A_n$, B, there *exists* a (Hilbert-type) deduction of B from $A_1,\ldots,A_n$. The significance of this latter result is that the Hilbert-type system of Section 2.6, which does not have the subformula property, is equivalent to the tableaux system of this section, which does have the subformula property. (This result is essentially based on the work of G. Gentzen, 1934-5.)

In order to show that our notions of tableau-deducibility (Def. 2.18) and (Hilbert-type) deducibility (Def. 2.11) are equivalent, we first prove the following.

Theorem 2.27. *(i) If B is tableau-deducible from $A_1,\ldots,A_n$, i.e., $A_1,\ldots,A_n \vdash' B$, then B is deducible from $A_1,\ldots,A_n$, i.e., $A_1,\ldots,A_n \vdash B$. In particular, for $n = 0$:*
(ii) If $\vdash' B$, then $\vdash B$.

Proof. Suppose $A_1,\ldots,A_n \vdash' B$, i.e., B is tableau-deducible from $A_1,\ldots,A_n$. It suffices to show:

for every sequent $S = \{TD_1,\ldots,TD_k,\ FE_1,\ldots,FE_m\}$ in a tableau-deduction of B from $A_1,\ldots,A_n$ it holds that $D_1,\ldots,D_k \vdash E_1 \vee \ldots \vee E_m$. (*)

Consequently, because $\{TA_1,\ldots,TA_n,\ FB\}$ is the first (upper) sequent in any given tableau-deduction of B from $A_1,\ldots,A_n$, we have that $A_1,\ldots,A_n \vdash B$.

The proof of (*) is tedious, but has a simple plan: the statement is true for the closed sequents in a tableau-deduction, and the statement remains true if we go up in the tableau-deduction via the T and F rules.

Basic step: Any closed sequent in a tableau-deduction of B from $A_1,\ldots,A_n$ is of the form $\{TD_1,\ldots,TD_k,\ TP,\ FP,\ FE_1,\ldots,FE_m\}$. So, we have to show that $D_1,\ldots,D_k,\ P \vdash P \vee E_1 \vee \ldots \vee E_m$. And this is straightforward: $D_1,\ldots,D_k,\ P \vdash P$ and $P \vdash P \vee E_1 \vee \ldots \vee E_m$.

Induction step: We have to show that for all rules the following is the case: if (*) holds for all lower sequent(s) in the rule (induction hypothesis), then (*) holds for the upper sequent in the rule. For convenience, we will suppose that $S = \{TD,\ FE\}$ in all rules.

Rule $T\wedge$:
$$\begin{array}{c} TD,\ FE,\ T\,B \wedge C \\ TD,\ FE,\ TB,\ TC \end{array}$$

Suppose $D,\ B,\ C \vdash E$ (induction hypothesis). To show: $D,\ B \wedge C \vdash E$. This follows immediately, because $B \wedge C \vdash B$ and $B \wedge C \vdash C$.

Rule $F\wedge$:
$$\begin{array}{c} TD,\ FE,\ F\,B \wedge C \\ TD,\ FE,\ FB \mid TD,\ FE,\ FC \end{array}$$

Suppose $D \vdash E \vee B$ and $D \vdash E \vee C$ (induction hypothesis). To show: $D \vdash E \vee (B \wedge C)$. It suffices to show that $E \vee B,\ E \vee C \vdash E \vee (B \wedge C)$. Now it is clear that $B,\ E \vdash E \vee (B \wedge C)$ and $B,\ C \vdash E \vee (B \wedge C)$. Hence, by $\vee$-elimination, $B,\ E \vee C \vdash E \vee (B \wedge C)$. But also $E,\ E \vee C \vdash E \vee (B \wedge C)$. Hence, again by $\vee$-elimination, $E \vee B,\ E \vee C \vdash E \vee (B \wedge C)$.

Rule $T\vee$: $TD,\ FE,\ T\ B\vee C$

$TD,\ FE,\ TB\ |\ TD,\ FE,\ TC$

Suppose $D,\ B\vdash E$ and $D,\ C\vdash E$ (induction hypothesis). To show: $D,\ B\vee C\vdash E$. This follows from the induction hypothesis by $\vee$-elimination.

Rule $F\vee$: $TD,\ FE,\ F\ B\vee C$

$TD,\ FE,\ FB,\ FC$

Suppose $D\vdash(E\vee B)\vee C$ (induction hypothesis). To show: $D\vdash E\vee(B\vee C)$. It suffices to show that $(E\vee B)\vee C\vdash E\vee(B\vee C)$.
It is clear that $E\vdash E\vee(B\vee C)$ and also $B\vdash E\vee(B\vee C)$. Hence, by $\vee$-elimination, $E\vee B\vdash E\vee(B\vee C)$. Since also $C\vdash E\vee(B\vee C)$, again by $\vee$-elimination, $(E\vee B)\vee C$ $\vdash E\vee(B\vee C)$.

Rule $T\rightarrow$: $TD,\ FE,\ T\ B\rightarrow C$

$TD,\ FE,\ FB\ |\ TD,\ FE,\ TC$

Suppose $D\vdash E\vee B$ and $D,\ C\vdash E$ (induction hypothesis). To show: $D,\ B\rightarrow C\vdash E$. By Exercise 2.50 $E\vee B,\ B\rightarrow C\vdash E\vee C$; hence, by the first induction hypothesis, $D,\ B\rightarrow C\vdash E\vee C$. (1)
From the second induction hypothesis, by the deduction theorem, $D\vdash C\rightarrow E$. (2)
By Exercise 2.50 $E\vee C,\ C\rightarrow E\vdash E\vee E$; hence, from (1) and (2): $D,\ B\rightarrow C\vdash E\vee E$. But by $\vee$-elimination $E\vee E\vdash E$. Hence $D,\ B\rightarrow C\vdash E$.

Rule $F\rightarrow$: $TD,\ FE,\ F\ B\rightarrow C$

$TD,\ FE,\ TB,\ FC$

Suppose $D,\ B\vdash E\vee C$ (induction hypothesis). To show: $D\vdash E\vee(B\rightarrow C)$.
From weak negation elimination, applying the deduction theorem, it follows that $\neg B\vdash B\rightarrow C$; hence $D,\ \neg B\vdash B\rightarrow C$. Hence $D,\ \neg B\vdash E\vee(B\rightarrow C)$. (1)
By Exercise 2.50 $E\vee C,\ C\rightarrow(B\rightarrow C)\vdash E\vee(B\rightarrow C)$. So, since $C\rightarrow(B\rightarrow C)$ is an axiom, it follows that $E\vee C\vdash E\vee(B\rightarrow C)$. So, by the induction hypothesis, $D,\ B\vdash E\vee(B\rightarrow C)$. (2)
From (1) and (2), by $\vee$-elimination $D,\ B\vee\neg B\vdash E\vee(B\rightarrow C)$. But, by Exercise 2.52, $\vdash B\vee\neg B$. Hence, $D\vdash E\vee(B\rightarrow C)$.

Rule $T\neg$: $TD,\ FE,\ T\ \neg B$

$TD,\ FE,\ FB$

Suppose $D\vdash E\vee B$ (induction hypothesis). To show: $D,\ \neg B\vdash E$. In order to do this, it suffices to prove that $E\vee B,\ \neg B\vdash E$.
By Exercise 2.53 $\neg B,\ \neg E\vdash\neg(E\vee B)$ and hence also $E\vee B,\ \neg B,\ \neg E\vdash\neg(E\vee B)$. But also $E\vee B,\ \neg B,\ \neg E\vdash E\vee B$. Hence, by $\neg$-introduction $E\vee B,\ \neg B\vdash\neg\neg E$. So, by double negation elimination $E\vee B,\ \neg B\vdash E$.

Rule $F\neg$: $TD,\ FE,\ F\ \neg B$

$TD,\ FE,\ TB$

Suppose $D,\ B\vdash E$ (induction hypothesis). To show: $D\vdash E\vee\neg B$.
From the induction hypothesis, $D,\ B\vdash E\vee\neg B$. (1)
From $\neg B\vdash E\vee\neg B$ it follows that $D,\ \neg B\vdash E\vee\neg B$. (2)
From (1) and (2) it follows by $\vee$-elimination that $D,B\vee\neg B\vdash E\vee\neg B$. By Exercise 2.52 $\vdash B\vee\neg B$ and hence $D\vdash E\vee\neg B$. □

With the help of tableaux we may give a constructive proof of the *interpolation theorem.*

Theorem 2.28 (Interpolation theorem for propositional logic). *Suppose $A \vdash' B$, $\nvdash' \neg A$ and $\nvdash' B$. Then there is a formula C such that every atomic formula that occurs in C also occurs in both A and B (so, C is in the joint vocabulary of A and B) and $A \vdash' C$ and $C \vdash' B$.*

Example 2.25. $(P \vee \neg Q) \wedge R \vdash' (Q \to P) \vee S$. Then for $C = P \vee \neg Q$, we have $(P \vee \neg Q) \wedge R \vdash' C$ and $C \vdash' (Q \to P) \vee S$.

Proof. Let A and B as mentioned in the interpolation theorem. Because $A \vdash' B$, any completed tableau starting with the initial sequent $\{TA,\ FB\}$ is closed, i.e., all its branches are closed. (*)
Since $\nvdash' \neg A$ we know that any completed tableau starting with $F\neg A$ (or, equivalently, TA) has at least one open (completed) branch $\mathscr{B}$. And since $\nvdash' B$, we know there any completed tableau starting with the initial sequent $\{FB\}$ has at least one open branch. Let $\mathscr{T}_A$ be a completed tableau starting with TA and $\mathscr{T}_B$ a completed tableau sarting with FB. We may assume that a tableau is closed if and only if it is atomically closed, i.e., every branch contains for some *atomic* formula P both TP and FP. For any open branch $\mathscr{B}$ in $\mathscr{T}_A$, we define the sets $\mathscr{B}^1$ and $\mathscr{B}^0$: $\mathscr{B}^1 = \{P \mid TP$ occurs in $\mathscr{B}$ and FP occurs in some open branch of $\mathscr{T}_B\}$ and $\mathscr{B}^0 = \{\neg P \mid FP$ occurs in $\mathscr{B}$ and TP occurs in some open branch of $\mathscr{T}_B\}$.

By (*) the union of $\mathscr{B}^0$ and $\mathscr{B}^1$ is non empty and so the following sentence is well-defined: $C(\mathscr{B})$:= the conjunction of all formulas in $\mathscr{B}^1 \cup \mathscr{B}^0$. Finally, the sentence C is defined as the disjunction of all formulas $C(\mathscr{B})$, where $\mathscr{B}$ is an open branch in the given tableau $\mathscr{T}_A$ starting with TA. Clearly, C is in the joint vocabulary of A and B. After some thinking it becomes clear that $A \vdash' C$ and $C \vdash' B$. □

Let us illustrate the proof for Example 2.25, where $A = \neg(Q \wedge \neg P) \wedge R$ and $B = (Q \to P) \vee S$. Let $\mathscr{T}_A$ be the following completed tableau starting with $F\ \neg(\neg(Q \wedge \neg P) \wedge R)$:

$$\begin{array}{c} F\ \neg(\neg(Q \wedge \neg P) \wedge R) \\ T\ \neg(Q \wedge \neg P) \wedge R \\ T\ \neg(Q \wedge \neg P),\ TR \\ F\ Q \wedge \neg P,\ TR \\ F\ \neg P,\ TR \mid FQ,\ TR \\ TP,\ TR \mid FQ,\ TR \end{array}$$

Both the left branch $\mathscr{B}_L$ and the right branch $\mathscr{B}_R$ of this tableau are open. Now, by definition, $\mathscr{B}_L^1 = \{P\}$, since there is an open branch starting with $F(Q \to P) \vee S$ that contains FP:

$$\begin{array}{c} F\ (Q \to P) \vee S \\ F\ (Q \to P),\ FS \\ TQ,\ FP,\ FS \end{array}$$

Note that $\mathscr{B}_L^0$ is empty. So, by definition, $C(\mathscr{B}_L) = P$.

By definition, $\mathscr{B}_R^1$ is empty and $\mathscr{B}_R^0 = \{\neg Q\}$, since there is an open branch starting with $F(Q \to P) \vee S$ that contains TQ. So, by definition, $C(\mathscr{B}_R) = \neg Q$. Finally, $C = C(\mathscr{B}_L) \vee C(\mathscr{B}_R) = P \vee \neg Q$.

Exercise 2.62. (a) Show, by using $\neg$-introduction, that $A \to B \vdash \neg(A \wedge \neg B)$.
(b) Show that $A \to B \vdash' \neg(A \wedge \neg B)$.
(c) Show that $A \to B \models \neg(A \wedge \neg B)$ by verifying that it is impossible that $A \to B$ is 1 and $\neg(A \wedge \neg B)$ is 0 in some line of the truth table. Note the analogy in (b) and (c).

Exercise 2.63. (a) Show, by using the deduction theorem three times, that $\vdash (A \to B) \to ((B \to C) \to (A \to C))$.
(b) Show that $\vdash' (A \to B) \to ((B \to C) \to (A \to C))$.
(c) Show that $\models (A \to B) \to ((B \to C) \to (A \to C))$ by verifying that it is impossible that this formula is 0 in some line of its truth table. Note the analogy in (b) and (c).

Exercise 2.64. Prove the following statements:
(a) $A \to B,\ \neg A \to B \vdash' B$
(b) $\neg B \to \neg A \vdash' A \to B$
(c) $\neg(A \wedge B) \vdash' \neg A \vee \neg B$
(d) $\neg(A \wedge \neg B) \vdash' A \to B$
(e) $A \to B \vdash' \neg A \vee B$
(f) $A \to B \vee C \vdash' (A \to B) \vee (A \to C)$

Exercise 2.65. a) Translate the following argument in the language of propositional logic.

If it rains [R], then John goes for a walk [W].
If it does not rain, then John makes a bicycle tour [B].
John does not make a bicycle tour.
Therefore: John goes for a walk.

b) Construct a tableau-deduction of the putative conclusion from the premisses or a counterexample (i.e., a line in the truth table in which all premisses are 1 and the putative conclusion is 0) from a failed attempt to do so.

Exercise 2.66. a) Translate the following argument in the language of propositional logic.

If it rains [R], then John does not go for a walk.
If John goes for a walk [W], then he is happy [H].
It does not rain.
Therefore: John is happy.

b) Construct a tableau-deduction of the putative conclusion from the premisses or a counterexample (i.e., a line in the truth table in which all premisses are 1 and the putative conclusion is 0) from a failed attempt to do so.

Exercise 2.67. (a) Verify that the (logical) axioms for (classical) propositional calculus of Section 2.6 are tableau-provable.
(b) Check that it is not a simple matter to prove: if $\vdash' A$ and $\vdash' A \to B$, then $\vdash' B$. Hence, the converse of Theorem 2.27, if $\vdash A$, then $\vdash' A$, to be shown in Section 2.9, is not a trivial result. However, one easily shows that $A,\ A \to B \vdash' B$ does hold.

Exercise 2.68. Show right from the definitions that
(a) if $\vdash' A$ or $\vdash' B$, then $\vdash' A \vee B$;
(b) if $\vdash' A \wedge B$, then $\vdash' A$ and $\vdash' B$.

Exercise 2.69. (a) Show that $\neg P,\ (P \to Q) \to P \vdash P$ by using weak negation elimination and the deduction theorem.
(b) Show that $P \vee \neg P,\ (P \to Q) \to P \vdash P$ by using (a) and $\vee$-elimination.
(c) Show that $\vdash ((P \to Q) \to P) \to P$ (Peirce's law) by using (b), Exercise 2.52. and the deduction theorem.
Compare the complexity of the proof of $\vdash ((P \to Q) \to P) \to P$ with the simplicity of the proof of $\vdash' ((P \to Q) \to P) \to P$. Note also that, although in Peirce's law implication is the only connective, we needed weak negation- and $\vee$-elimination in order to show that Peirce's law is (logically) provable (see Exercise 2.44).

2.9 Completeness of classical propositional logic

So far we have established the following results; for convenience, we use the Greek letter Γ to indicate a (possibly infinite) collection of formulas.
Theorem 2.27: if $\Gamma \vdash' B$, then $\Gamma \vdash B$.
Theorem 2.20: if $\Gamma \vdash B$, then $\Gamma \models B$ (soundness). In this section we shall prove *completeness*, i.e., every valid consequence of given premisses Γ can be (logically) deduced from Γ: if $\Gamma \models B$, then $\Gamma \vdash' B$.

This shows that the three notions $\Gamma \vdash' B$ (B is tableau-deducible from Γ), $\Gamma \vdash B$ (B is deducible from Γ) and $\Gamma \models B$ (B is a valid consequence of Γ) are equivalent.

The intuitive 'B is a logical consequence of the premisses in Γ' (without reference to the structure of the atomic formulas in B and Γ) has been made mathematically precise in three different ways: $\Gamma \vdash' B$, $\Gamma \vdash B$ and $\Gamma \models B$. Since these three mathematical notions, although intensionally quite different, turn out to be equivalent, we may say (after the results we are about to prove) that *we indeed have captured in a mathematically definite sense the intuitive notion of 'B is a logical conclusion from Γ'*. (See also the discussion following Theorem 2.21.)

In proving the completeness of classical propositional logic, a procedure of searching for a tableau-deduction of B from given premisses $A_1, \dots, A_n$ is presented, which will end after finitely many steps and then either gives such a deduction or shows that such a deduction cannot exist. This algorithm thus yields a decision procedure for the (classical) propositional logic. This shall provide us an opportunity to dwell upon automated theorem proving.

Given formulas B and $A_1, \dots, A_n$, the tableaux rules suggest a *procedure of searching for a tableau-deduction* of B from $A_1, \dots, A_n$:
start with $TA_1, \dots, TA_n,\ FB$ and apply all the appropriate rules in some definite fixed order, the choice of ordering being unimportant (at least, if we do not care about efficiency); in an application of rule $T\to$ to, for example, $S,\ T\ P \to Q$ we make two branches, one with $S,\ FP$ and the other with $S,\ TQ$ and similarly for applications of the rules $F\wedge$ and $T\vee$.

Example 2.26. 1) The tableau starting with $F\ (P \to Q) \to (Q \to \neg P)$ is composed of the following two branches:

$F\ (P \to Q) \to (Q \to \neg P)$	and $F(P \to Q) \to (Q \to \neg P)$
$T\ P \to Q,\ F\ Q \to \neg P$	$T\ P \to Q,\ F\ Q \to \neg P$
$FP,\ F\ Q \to \neg P$	$TQ,\ F\ Q \to \neg P$
$FP,\ TQ,\ F\ \neg P$	$TQ,\ TQ,\ F\ \neg P$
$FP,\ TQ,\ TP$	$TQ,\ TQ,\ TP$

The first branch for $(P \to Q) \to (Q \to \neg P)$ is closed; the second one is completed and open. Note that if we assign the value 1 to both P and Q, corresponding with the fact that both TP and TQ occur in the open branch, the formula $(P \to Q) \to (Q \to \neg P)$ is assigned the value 0, corresponding with the fact that $F\ (P \to Q) \to (Q \to \neg P)$ occurs in the open branch. We shall see in Lemma 2.2 that this is not accidental.

2) The tableau starting with $T\ P \to Q,\ F\ \neg Q \to \neg P$ is composed of the following two branches:

$T\ P \to Q,\ F\ \neg Q \to \neg P$	and $TP \to Q,\ F\ \neg Q \to \neg P$
$FP,\ F\ \neg Q \to \neg P$	$TQ,\ F\ \neg Q \to \neg P$
$FP,\ T\ \neg Q,\ F\ \neg P$	$TQ,\ T\ \neg Q,\ F\ \neg P$
$FP,\ FQ,\ F\ \neg P$	$TQ,\ FQ,\ F\ \neg P$
$FP,\ FQ,\ TP$	$TQ,\ FQ,\ TP$

Both branches starting with $T\ P \to Q,\ F\ \neg Q \to \neg P$ are closed. Note that the two branches together yield a tableau-deduction of $\neg Q \to \neg P$ from $P \to Q$, just as a tableau-proof of $(P \to Q) \to (\neg Q \to \neg P)$. The correctness of this statement is not accidental either and follows immediately from the definition of a tableau-deduction and the structure of our procedure of searching for a tableau-deduction; see Lemma 2.3.

Definition 2.19. Let τ be a completed tableau branch which is open. Then i_τ is the interpretation defined by $i_\tau(P) = 1$ if TP occurs in τ, $i_\tau(P) = 0$ if TP does not occur in τ.

Lemma 2.2. *Let τ be a completed tableau branch which is open. Then for each formula E: a) if TE occurs in τ, then $i_\tau(E) = 1$, and*
b) if FE occurs in τ, then $i_\tau(E) = 0$.

Proof. The proof is by induction on the construction of E. Let τ be a completed tableau branch which is open.
Basic step. If $E = P$ (atomic formula) and TP occurs in τ, then by definition $i_\tau(P) = 1$. If $E = P$ and FP occurs in τ, then - since τ is open - TP does not occur in τ and hence by definition $i_\tau(P) = 0$.
Induction step. Suppose that a) and b) have been shown for C and D (induction hypothesis). We want to prove a) and b) for $C \wedge D$, $C \vee D$, $C \to D$ and $\neg C$.

If $E = C \wedge D$ and $T\ C \wedge D$ occurs in τ, then - because τ is completed - both TC and TD occur in τ. Hence, by the induction hypothesis, $i_\tau(C) = 1$ and $i_\tau(D) = 1$. So, $i_\tau(C \wedge D) = 1$.
If $E = C \wedge D$ and $F\ C \wedge D$ occurs in τ, then - because τ is completed - FC occurs in τ or FD occurs in τ. Hence, by the induction hypothesis, $i_\tau(C) = 0$ or $i_\tau(D) = 0$. So, $i_\tau(C \wedge D) = 0$.

The other cases, $E = C \vee D$, $E = C \to D$ and $E = \neg C$, are treated similarly. □

Lemma 2.3. *If all branches in a tableau with initial sequent* $\{TA_1,\ldots,TA_n,FB\}$ *are closed, then* $A_1,\ldots,A_n \vdash' B$.

Proof. This follows from the definition of a tableau with $\{TA_1,\ldots,TA_n,FB\}$ as initial sequent and from the observation that there are only finitely many different branches in such a tableau. □

Lemma 2.2 and 2.3 together yield the completeness theorem.

Theorem 2.29 (completeness of classical propositional logic).
a) If $A_1,\ldots,A_n \models B$*, then* $A_1,\ldots,A \vdash' B$*. In particular, if* $n = 0$*:*
b) If $\models B$*, then* $\vdash' B$.

Proof. Suppose $A_1,\ldots,A_n \models B$. Apply the procedure of searching for a tableau-deduction of B from $A_1,\ldots,A_n$. If there were a completed tableau branch τ starting with $TA_1,\ldots,TA_n,FB$ which is open, then by Lemma 2.2, because $TA_1,\ldots,TA_n$ and FB occur in such a τ, $i_\tau(A_1) = \ldots = i_\tau(A_n) = 1$ and $i_\tau(B) = 0$. This would contradict that $A_1,\ldots,A_n \models B$. Hence, all tableau branches starting with $TA_1,\ldots,TA_n$, FB are closed. So, by Lemma 2.3, $A_1,\ldots,A_n \vdash' B$. □

Remark 2.2. Our procedure of searching for a tableau-deduction of B from given premisses $A_1,\ldots,A_n$ will end after finitely many steps and then either give a tableau-deduction of B from $A_1,\ldots,A_n$, indicating that $A_1,\ldots,A_n \vdash' B$, or an interpretation i such that $i(A_1) = \ldots = i(A_n) = 1$ and $i(B) = 0$, indicating that $A_1,\ldots,A_n \not\models B$.

Corollary 2.4 (Decidability of classical propositional logic). *Classical propositional logic is decidable, i.e., we have an effective method (algorithm) to decide, given any finite set of formulas* B*,* $A_1,\ldots,A_n$*, whether* B *is tableau-deducible from* $A_1,\ldots,A_n$ *or not.*

Note that in Section 2.3 we have already given an effective method (algorithm) to decide whether or not B is a valid consequence of $A_1,\ldots,A_n$ for any finite set of formulas $A_1,\ldots,A_n$.

The tableaux system for classical propositional calculus can easily be modified and/or completed to a tableaux system for intuitionistic logic and for many intensional (modal) logics. In all cases the completeness proof given above can be adapted to a completeness proof for the logic in question. This type of proof has an advantage over some other completeness proofs in that it is *constructive*.

Automated theorem proving In the case of the classical propositional calculus an effective method has been given above to decide, given any finite set of formulas B, $A_1,\ldots,A_n$, whether B is tableau-deducible from $A_1,\ldots,A_n$ or not. This algorithm can be formulated in an appropriate programming language such as Prolog (see, for instance, Kogel-Ophelders [17]) and then a computer, when provided with formulas $B, A_1,\ldots,A_n$, is able to compute whether B is a theorem on the basis of the hypotheses $A_1,\ldots,A_n$ or not.

So, a computer, provided with the appropriate software, is able to simulate reasoning and in that case one may say that it disposes of *Artificial Intelligence*. By adding to such a computer-program a number of data, $A_1,\ldots,A_n$, concerning a small and well-described subject, the so-called *knowledge base*, the computer is able to draw conclusions from those data. If $A_1,\ldots,A_n$ represent someone's expertise, one speaks of an *expert system*. And if the knowledge base consists of Euclid's axioms for geometry or Peano's axioms for number theory or of axioms for some other part of mathematics, one speaks of *automated theorem proving*.

So the basic ideas underlying expert-systems and automated theorem proving are very simple. However, in practice there may be a lot of complications. Without being exhaustive let us mention some of them.

1. The language of propositional logic may be too restrictive. For instance, in Chapter 1 we have already seen that the argument

All men are mortal.
Socrates is a man.
Therefore, Socrates is mortal.

cannot be adequately formulated in the propositional language. For that reason the propositional language will be extended to the predicate language in Chapter 4.

2. However, if one adapts the construction of a completed tableau with initial branch $\{TA_1,\ldots,TA_n,FB\}$ to the case that B, $A_1,\ldots,A_n$ are formulas of the predicate language, this construction no longer yields a decision: if no logical deduction exists, the tableau construction may continue forever, without ever knowing that this construction will come to an end; so, in this case the tableau construction may not stop. For more details see Subsection 4.4.2.

3. Even in the case of the propositional language, the time and space needed to search for a logical deduction of B from $A_1,\ldots,A_n$ may grow very fast in the event n is big or B, $A_1,\ldots,A_n$ are (very) complex; see Subsection 2.3.1.

4. If the knowledge base consists of Peano's axioms for number theory (see Section 5), this knowledge base contains the axiom schema of induction, and hence infinitely many axioms. Searching for a logical deduction of a given formula B from infinitely many axioms requires a *strategy*, without which such a search is hopeless.

5. If the knowledge base consists of someone's expertise, it may contain uncertain and/or incomplete information. For instance, it may be likely, but uncertain, that there is oil in the ground. An expert-system may have to deal with uncertain knowledge and then its conclusions will have a certain degree of probability, which has to be computed. This is a far from trivial matter. Also the information in the knowledge base may be incomplete in order to be able to draw a certain conclusion.

6. Building an expert-system is more than just providing an inference mechanism: the system should also be able to explain *how* the conclusion was established or *why* the conclusion cannot be drawn.

2.10 Paradoxes; Historical and Philosophical Remarks

2.10.1 Paradoxes

Paradoxes have been important for making progress in science and philosophy. In what follows a number of statements of the type $B \rightleftarrows \neg B$ are presented. Because statements of this type cannot possibly be true, in other words are inconsistent, these results are known as *paradoxes*. The reader easily checks the following theorem:

Theorem 2.30. *a) For each formula B,* $\models \neg(B \rightleftarrows \neg B)$.
b) If $A_1, \ldots, A_n \models B \rightleftarrows \neg B$*, then* $\models \neg(A_1 \wedge \ldots \wedge A_n)$.

So, if for some formula B, $B \rightleftarrows \neg B$ is a valid consequence of hypotheses $A_1, \ldots, A_n$, then at least one of the hypotheses must be false. In practice, the problem frequently is that we are not aware of the hypotheses we are using in deriving a paradox.

In his paper *'Paradox'*, W.V. Quine [21] distinguishes three types of paradox: antinomies, veridical and falsidical paradoxes. Below we shall discuss these three types and consider examples of each of them.

Antinomies There is the old *paradox of the liar*: A man says that he is lying. If he speaks the truth, he is lying. And if he is lying, he speaks the truth. Hence, he speaks the truth if and only if he does not.

A more recent version of this paradox is the one of A. Tarski [24] in his *'Truth and Proof'*. Consider the following sentence.

s: The underlined sentence is false

Here *s* is just an abbreviation for: the underlined sentence is false. But what is the object the name 'the underlined sentence' refers to? Up till now there is no underlined sentence. By underlining sentence *s*, we achieve that sentence *s* says of itself that it is false, just as the man in the paradox says of himself that he is lying.

s: <u>The underlined sentence is false</u>

When one refers to an object, one usually uses a name for that object. One and the same object may have different names. For instance, 'Harrie de Swart' and 'the author of this book' are two different names for the same person. Usually, when referring to a sentence or, more generally, a linguistic object, one may form its name by putting the sentence in question between quotation marks. But another name for that same sentence may be formed by underlining the sentence in question, after which 'the underlined sentence' is another name for the same sentence. So, having underlined sentence *s*, *s* has (at least) the following two names: '*s*'; the underlined sentence. Consequently, by replacing one name by another one:

(1) '*s*' is false if and only if the underlined sentence is false.

On the other hand we have the *principle of adequacy*: for each sentence *p*, '*p*' is true if and only if *p*; where '*p*' is again a name for the sentence *p*. For example, 'snow is white' is true if and only if snow is white. Now using this principle of adequacy, we find

's' is true if and only if s, i.e.,
(2) 's' is true if and only if the underlined sentence is false.

(1) and (2) together yield: 's' is false if and only if 's' is true.

The paradox of the liar, in one form or another, is a special kind of paradox, an *antinomy*: an absurd statement, that cannot be true, with a correct argument, and whose premisses are not in themselves absurd. However, if $B \rightleftarrows \neg B$ is a valid consequence of premisses $A_1, \ldots, A_n$, we know we have to revise our premisses. It is typical of an antinomy that we are very surprised that such a revision is necessary, because the premisses accepted seem more than plausible and seem completely in accordance with our intuition. In order to be able to 'solve' an antinomy, a major revision in our way of thinking is necessary. Because everything we do in the derivation of an antinomy seems so natural and evident, we are generally not very conscious of what precisely our premisses are.

Through all ages the antinomies have caused concern to philosophers. According to a foolish tradition preserved by Diogenes Laertius, Diodorus Cronus (ca. 300 B.C.) committed suicide because he was not immediately able to solve the logical puzzle posed by the paradox of the liar. (See W & M Kneale [16], p. 113.)

In his paper *Truth and Proof*, A. Tarski [24] argues that the paradox of the liar forces us to give up our silent assumption that object language and meta language do not have to be distinguished. But when we say that a sentence 's' is true, we are saying something *about* sentence s. If s belongs to a language L_0, the sentence ' 's' is true' is a statement *about* a sentence of L_0 and hence a statement in the metalanguage L_1 of L_0. If we take care to distinguish predicates true_0, true_1, false_0, false_1, and so on, for the truth/falsity predicates in the different languages, the paradox of the liar disappears:

Again, let s be an abbreviation for: the underlined sentence is true_0. Next, let us underline this sentence:

s: The underlined sentence is false_0

Then, again replacing one name by another one:

(1a) 's' is false_1 if and only if the underlined sentence is false_1.

And by the principle of adequacy

(2a) 's' is true_1 if and only if the underlined sentence is false_0.

And now (1a) and (2a) are no longer contradictory!

If we wish to avoid contradictions, we must insist that what we ordinarily call English is in reality an infinite sequence L_0, L_1, $L_2, \ldots$ of languages, in which L_{n+1} is a metalanguage in relation to L_n.

Another way to escape the antinomy of the liar is by introducing a technical restriction on the class of sentences regarded as possessing a truth value. According to Ryle [22], sentences of the form 'the such-and-such sentence is false' should not be regarded as having a truth value unless it is possible to attach a 'namely-rider'. For instance, in 'the first thing that Plato said to Aristotle is true' we can insert a clause, 'the first thing that Plato said to Aristotle, namely ..., is true', which may alter its meaning, but does not alter its truth-value. But in the paradoxical 'the underlined sentence is false', if we try to insert such a clause, 'the underlined sentence, namely 'the underlined sentence is false', is false' we get a new description (indirect) of a sentence which must again be supplied with a namely-rider. As this process never ends, the original sentence has no truth value, whereas in the Plato example, we get down to something of the form '...' is true, where the quoted part does not involve the notions of truth and falsehood.

The paradox of the liar is an antinomy at the level of sentences. At the level of subjects and singular descriptions there is the *antinomy of Berry*, to be discussed in Chapter 4. And at the level of predicates there is the *antinomy of Russell*, better known as *Russell's paradox*, which will be discussed in Chapter 3.

Besides antinomies, like those of the liar, of Berry and of Russell, W.V. Quine also distinguishes other, less serious, paradoxes: veridical and falsidical paradoxes.

Veridical paradoxes A *veridical* or *truth-telling paradox* is a paradoxical statement that on reflection turns out to yield a somewhat astonishing, but true, proposition.

Example 2.27. 1. Frederic has reached the age of twenty-one without having more than five birthdays.
2. The **barber paradox**: In a certain village there is a barber who shaves precisely those men in the village who do not shave themselves. Question: does the barber shave himself? Each man in the village is shaved by the barber if and only if he does not shave himself. Hence, in particular, the barber shaves himself if and only if he does not shave himself.

Both paradoxes are alike in the sense that at first sight they seem to prove absurdities by decisive arguments. The Frederic-paradox is a truth-telling paradox if we conceive the statement as the abstract truth that one can be $4n$ ($n = 0, 1, 2, \ldots$) years old at one's n^{th} birthday, namely if one has been born on February 29. The barber-paradox contains a reductio ad absurdum: from the, not explicitly mentioned, premiss that such a barber exists, we derive an absurdity of the form $B \rightleftarrows \neg B$. Hence the assumption is false and no village can have a barber who shaves all and only those men in the village who do not shave themselves.

The difference between an antinomy and a veridical paradox is that in the latter case we are only slightly astonished that we have to give up one of the premisses like the existence of a village-barber as described above, while in the case of an antinomy we are forced to give up very fundamental ideas and a major revision in our way of thinking is needed.

Falsidical paradoxes A *falsidical paradox* is a paradoxical statement that really is false, the argument backing it up containing some impossible hidden assumption or involving a fallacy. Typical examples of falsidical paradoxes are:

Example 2.28. 1. The comic mis-proof that $2 = 1$: Let $x = 1$. Then $x^2 = x$. Hence $x^2 - 1 = x - 1$. Dividing both sides by $x - 1$, we conclude that $x + 1 = 1$. Hence, because $x = 1$, $2 = 1$.

2. Three men agree to share a hotel room overnight, splitting the charge of \$ 30 three ways, with each man paying \$ 10. After they have gone to their room, the clerk realizes he should only have charged them \$ 25 and sends the bellboy up with \$ 5 to be returned to them. The bellboy, realizing how hard it will be to make change, pockets \$ 2 and returns \$ 1 to each man. Thus the men have each paid \$ 9, for a total of \$ 27 and the bellboy has \$ 2, for a total of \$ 29. One dollar of the original thirty is missing.
3. Zeno's paradox of Archilles and the Tortoise.

```
 A
 | - - - - - - - - - - | o o o o o o o |. . . . |
0                        T
                      | - - - - - - - |o o o o |. . . . |
                      1               1.1       1.11      1.111
```

Suppose A(rchilles) and the T(ortoise) start to run at the same time and A runs 10 times as fast as T does. Suppose also that in the starting position A is in position 0, one mile behind T, which hence is in position 1. While A runs from 0 to the starting position of T, T covers a distance of 0.1 mile since its velocity has been supposed to be $\frac{1}{10}$ of that of A. And while A runs from position 1 to position 1.1, T covers a distance of 0.01 mile, thus arriving at position 1.11. And while A runs from position 1.1 to position 1.11, T runs from position 1.11 to position 1.111. And so on. Consequently, A will never pass T.

In a falsidical paradox there is always a fallacy or some impossible hidden assumption in the argument and in addition the statement must look absurd and be false.

In the 'proof' of $2 = 1$ we divided by $x - 1$, which is 0 because x was supposed to be 1. In the hotel paradox the number 2 is added wrongly to 27: 2 should be subtracted from 27 in order to determine the price, 25 dollars, of the hotel room.

In the case of Archilles and the Tortoise the impossible hidden assumption is that the infinite process of Archilles running to the position where the tortoise was a moment ago, lasts infinitely long. In fact, however, if Archilles needs 0.1 hour for one mile, the infinite process will last only $0.1 + 0.01 + 0.001 + \ldots = 0.111\ldots = \frac{1}{9}$ hour, which is less than 0.12 hours. Within this time Archilles and the Tortoise will arrive at the same position and Archilles will pass the Tortoise. The process of Archilles passing the Tortoise may be thought of as consisting of infinitely many steps, but this infinite process is actually completed in $\frac{1}{9}$ hour (6 minutes and 40 seconds).

Only the antinomies cause a crisis of thought. Only an antinomy produces a self-contradiction via accepted means of reasoning. Only an antinomy requires that some tacitly accepted and trusted patterns of reasoning be made explicit and henceforth be avoided or revised.

The falsidical paradox of Zeno must have been a real antinomy in his day. It was thought as evident that a process consisting of infinitely many steps would last infinitely long. It is only because of the mathematical achievements of the 18th and 19th century that we know that some infinite sums, for example, $0.1 + 0.01 + 0.001 + \ldots = 0.111\ldots = \frac{1}{9}$ and $\frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots = 1$, are finite, while others, for example, $\frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \ldots$, are not. What is an antinomy for the one is a falsidical paradox for the other, given a lapse of a couple of thousands of years.

In the case of the paradox of Archilles and the Tortoise one should realize that points of space and time do not occur in our perception, but are mathematical idealizations. Points of space and time belong to the language of mathematics, not to the language of our perception. If we talk about Archilles passing the infinitely many points (positions) the Tortoise was a moment ago, we are speaking in terms of our mathematical model and not in terms of what we perceive.

Exercise 2.70. Is the following paradox an antinomy, a veridical or a falsidical one? A judge tells a condemned prisoner that he will be hanged either on Monday, Tuesday, Wednesday, Thursday or Friday of the next week, but that the day of the hanging will come as a surprise: he will not know until the last moment that he is going to be hanged on that day. The prisoner reasons that if the first four days go by without the hanging, he will *know* on Friday, that he is due to be hanged that day. So it cannot be on Friday that he will be hanged. But now with Friday eliminated, if the first three days go by without the hanging, he will know on Thursday that he is due to be hanged that day, and it would not be a surprise. So it cannot be Thursday. In the same way he rules out Wednesday, Tuesday and Monday, and convinces himself that he cannot be hanged at all. But he is very surprised on Wednesday when the executioner arrives at his cell. (See also Exercise 6.12 and its solution.)

Exercise 2.71. Is the following paradox an antinomy, a veridical or a falsidical one? A crocodile seizes a baby, and tells the mother that he will return it if the next thing she says to him is the truth, but will eat it if the next thing she says is false. The mother says 'you will eat the baby'. The crocodile will eat the baby if and only if he will let it go.

Exercise 2.72. (From S.C. Kleene [15], p. 40) The following riddle also turns upon the paradox of the liar. A traveller has fallen among cannibals. They offer him the opportunity to make a statement, attaching the conditions that if his statement be true, he will be boiled, and if it be false, he will be roasted. What statement should he make? (A form of this riddle occurs in Cervantes' "Don Quixote" (1605), II, 51.)

Exercise 2.73. (From S.C. Kleene [15], p. 37, 38) Every municipality in Holland must have a mayor, and no two may have the same mayor. Sometimes the mayor is a non-resident of the municipality. Suppose a law is passed setting aside a special area S exclusively for such non-resident mayors, and compelling all non-resident mayors to reside there. Suppose further that there are so many non-resident mayors that S has to be constituted a municipality. Where shall the mayor of S reside? (Mannoury, cf. van Dantzig [5])

Exercise 2.74. (From S.C. Kleene [15], p. 38) Suppose the Librarian of Congress compiles, for inclusion in the Library of Congress, a bibliography of all those bibliographies in the Library of Congress which do not list themselves. (Gonseth 1933) Should that bibliography list itself?

Exercise 2.75. From *Attic Nights* by Aulus Gellius, Book V, x:

> Among fallacious arguments the one which the Greeks call *ἀντιστρέφον* seems to be by far the most fallacious. Some of our own philosophers have rather appropriately termed such arguments *reciproca*, or 'convertible'. The fallacy arises from the fact that the argument that is presented may be turned in the opposite direction and used against the one who has offered it, and is equally strong for both sides of the question. An example is the well-known argument which Protagoras, the keenest of all sophists, is said to have used against his pupil Euathlus.
>
> For a dispute arose between them and an altercation as to the fee which had been agreed upon, as follows: Euathlus, a wealthy young man, was desirous of instruction in oratory and the pleading of causes. He became a pupil of Protagoras and promised to pay him a large sum of money, as much as Protagoras had demanded. He paid half of the amount at once, before beginning his lessons, and agreed to pay the remaining half on the day when he first pleaded before jurors and won his case. Afterwards, when he had been for some little time a pupil and follower of Protagoras, and had in fact made considerable progress in the study of oratory, he nevertheless did not undertake any cases. And when the time was already getting long, and he seemed to be acting thus in order not to pay the rest of the fee, Protagoras formed what seemed to him at the time a wily scheme; he determined to demand his pay according to the contract, and brought suit against Euathlus.
>
> And when they had appeared before the jurors to bring forward and to contest the case, Protagoras began as follows: 'Let me tell you, most foolish of youths, that in either event you will have to pay what I am demanding, whether judgement be pronounced for or against you. For if the case goes against you, the money will be due me in accordance with the verdict, because I have won; but if the decision be in your favour, the money will be due me according to our contract, since you will have won a case.'
>
> To this Euathlus replied: 'I might have met this sophism of yours, tricky as it is, by not pleading my own cause but employing another as my advocate. But I take greater satisfaction in a victory in which I defeat you, not only in the suit, but also in this argument of yours. So let me tell you in turn, wisest of masters, that in either event I shall not have to pay what you demand, whether judgement be pronounced for or against me. For if the jurors decide in my favour, according to their verdict nothing will be due you, because I have won; but if they give judgement against me, by the terms of our contract I shall owe you nothing, because I have not won a case.'
>
> Then the jurors, thinking that the plea on both sides was uncertain and insoluble, for fear that their decision, for whichever side it was rendered, might annul itself, left the matter undecided and postponed the case to a distant day. Thus a celebrated master of oratory was refuted by his youthful pupil with his own argument, and his cleverly devised sophism failed. [From the English translation by John C. Rolfe of *The Attic Nights of Aulus Gellius*, Book V, section X. Reprinted, Cambridge, Mass., 1967. The Loeb Classical Library, 195, pp. 404-409.]

2.10.2 Historical and Philosophical Remarks

Stoic Logic Aristotle is generally seen as the founding father of logic. Only at the beginning of the 20th century it became clear, among others by the work of the Polish logician Łukasiewicz, that in fact the Stoics (± 300 B.C.) developed a kind of propositional logic, while the logic of Aristotle is a small part of what we now call predicate logic, to be studied in Chapter 4. A typical inference-schema of the

Stoics runs as follows:

If the first, then the second.
The first.
Therefore, the second.

As a concrete example of this type of inference, they were accustomed to give:

If it is day, then it is light.
It is day.
Therefore, it is light.

A typical Aristotelian syllogism is: If all things with the predicate (property) P also satisfy the predicate Q, and all things with the predicate Q also satisfy the predicate R, then all things with the predicate P also satisfy the predicate R. A concrete instance of this would be: If all birds are animals and all animals are mortal, then also all birds are mortal.

As pointed out by Łukasiewicz, the Stoics were discussing the truth conditions for implication. The truth-functional account, as in our truth table for $\rightarrow$, is first known to have been proposed by Philo of Megara ca. 300 B.C. in opposition to the view of his teacher Diodorus Cronus. We know of this through the writings of Sextus Empiricus some 500 years later, the earlier documents having been lost. According to Sextus,

> Philo says that a sound conditional is one that does not begin with a truth and end with a falsehood. ... But Diodorus says it is one that neither could nor can begin with a truth and end with a falsehood. [Kneale, [16], p. 128]

There can be no doubt that what Sextus refers to is precisely the truth-functional connective that we have symbolized by the $\rightarrow$, for he says elsewhere,

> So according to him there are three ways in which a conditional may be true, and one in which it may be false. For a conditional is true when it begins with a truth and ends with a truth, like 'if it is day, it is light'; and true also when it begins with a falsehood and ends with a falsehood, like 'If the earth flies, the earth has wings'; and similarly a conditional which begins with a falsehood and ends with a truth is itself true, like 'If the earth flies, the earth exists'. A conditional is false only when it begins with a truth and ends with a falsehood, like 'If it is day, it is night'. [Kneale [16], p. 130]

So Sextus reports Philo as attributing truth values to conditionals just as in our truth table for $\rightarrow$, except for the order in which he lists the cases. Diodorus probably had in mind what later was called *strict implication*; see Chapter 6.

One of the Stoic principles noted by Łukasiewicz is as follows: an argument is valid if and only if the conditional proposition having the conjunction of the premisses as antecedent and the conclusion as consequent is logically true. The similarity of this principle to our Theorem 2.4 is obvious.

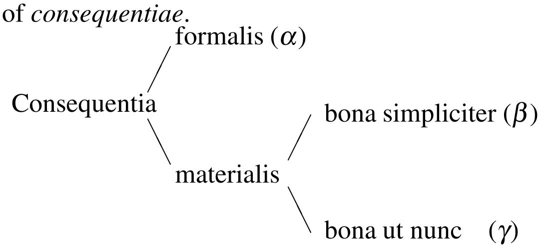
According to the Stoics, there were five basic types of undemonstrated, i.e., self evident, argument:

1. If the first, then the second; but the first. Therefore, the second.
2. If the first, then the second; but not the second. Therefore not the first.
3. Not both the first and the second; but the first. Therefore not the second.
4. Either the first or the second; but the first. Therefore not the second.
5. Either the first or the second; but not the second. Therefore the first.

These arguments are basic, it was maintained, in the sense that every valid argument can be reduced to them. Sextus Empiricus gives us two very clear examples of the analysis of an argument into its component basic arguments:
6. If the first, then if the first then the second; but the first. Therefore the second. (Composition of two type 1 undemonstrated arguments.)
7. If the first and the second, then the third; but not the third; on the other hand the first. Therefore not the second. (Composition of a type 2 and a type 3 undemonstrated argument.)
One of the theorems attributed to Chrysippus is:
8. Either the first or the second or the third; but not the first; and not the second. Therefore the third. (Composition of two type 5 undemonstrated arguments.)

Chrysippus himself is reported to have said that even dogs make use of this sort of argument. For when a dog is chasing some animal and comes to the junction of three roads, if he sniffs first at the two roads down which the animal did not run, he will rush off down the third road without stopping to smell. [See B. Mates [19], pp. 67-82 and W. & M. Kneale [16], pp. 158-176.]

Consequentiae In the Middle Ages several treatises on *consequentiae* were written. One of the more interesting ones is *In Universam Logicam Quaestiones*, formerly attributed to John Duns the Scot (1266-1308), but later to a Pseudo-Scot (? John of Cornwall). As we learn from Kneale [16], pp. 278-280, the Pseudo-Scot distinguishes various kinds of *consequentiae*.



Examples:

(α) Socrates currit et Socrates est albus, igitur album currit.
Socrates walks and Socrates is white, so something white walks.
(β) Homo currit, igitur animal currit.
A man walks, therefore a living being walks.
(γ) Socrates currit, igitur album currit.
Socrates walks, therefore something white walks.

Consequentiae formales are inferences made exclusively on the basis of the forms of the expressions involved. In *consequentiae materiales* the meaning of the premisses and the conclusion also has to be taken into account. But *consequentiae materiales* can always be reduced to consequentiae formales by making explicit the silently assumed premises. For instance, 'Socrates currit, igitur album currit' (Socrates walks, so something white walks) can be reduced to 'Socrates currit et Socrates est albus, igitur album currit' (Socrates walks and Socrates is white, so something white walks). The *Consequentiae materiales bona simpliciter* are those inferences in which the silently assumed premisses are necessary, like, for instance, ‘omnis homo est animal’ (every man is a living being). When the silently assumed premisses are contingent (not necessary), like, for instance, ‘Socrates est albus’ (Socrates is white), the Pseudo-Scot speaks of *consequentiae materiales bona ut nunc*.

Because of their amusing character, we present below two theorems and their proofs, as given by the Pseudo-Scot.

1. *Ad quamlibet propositionem implicantem contradictionem de forma sequitur quaelibet alia propositio in consequentia formali* (From a proposition which implies a formal contradiction, any proposition follows as a ‘consequentia formalis’).
2. *Ad quamlibet propositionem impossibilem sequitur quaelibet alia propositio non consequentia formali sed consequentia materiali bona simpliciter* (From a proposition which is impossible, any proposition follows not as a ‘consequentia formalis’ but as a ‘consequentia materialis bona simpliciter’).

Kneale [16], pp. 281-282, gives the following reconstruction of the proof of 1.

$$\dfrac{\dfrac{\begin{array}{c}\text{Socrates exists and}\\ \text{Socrates does not exist}\end{array}}{\text{Socrates does not exist}} \qquad \dfrac{\dfrac{\text{Socrates exists and Socrates does not exist}}{\text{Socrates exists}}}{\text{Socrates exists or a man is an ass}}}{\text{a man is an ass}}$$

And the Pseudo-Scot gives the following two proofs of 2:

1. Using 1., the *consequentia* ‘A man is an ass and a man is not an ass, therefore you are in Rome’ is formally valid. Since it is impossible that a man is an ass, it is necessary that a man is not an ass. And the Pseudo-Scot concludes that the *consequentia materialis* ‘A man is an ass, therefore you are at Rome’ is bona simpliciter, being reducible to a formally valid *consequentia* by addition of a necessarily true premise.
2. Supposing that ‘A man is not an ass’ is necessarily true, the Pseudo-Scot also gives the following derivation.

$$\dfrac{\dfrac{\text{A man is an ass}}{\text{A man is an ass or you are at Rome}} \qquad \text{A man is not an ass}}{\text{you are at Rome}}$$

Suggested reading on Medieval Logic: W. & M. Kneale, *The Development of Logic*; L.M. de Rijk, *Logica Modernorum*; P. Boehner, *Medieval Logic*; E. Moody, *Truth and Consequence in Medieval Logic*.

Frege’s Begriffsschrift (1879) Although an algebra of logic was initiated by Boole in 1847 and De Morgan in that same year, the propositional logic properly appeared with Frege’s *Begriffsschrift* in 1879, and in Russell’s work, especially in the *Principia Mathematica* by Whitehead and Russell, 1910-13.

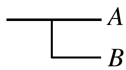
> The imprecision and ambiguity of ordinary language led Frege (1848-1925) to look for a more appropriate tool; he devised a new mode of expression, a language that deals with the 'conceptual content' and that he came to call 'Begriffsschrift'. This ideography is a 'formula language', that is, a *lingua characterica*, a language written with special symbols, 'for pure thought', that is, free from rhetorical embellishments, ... [Heijenoort [12], p. 1]

In the preface to his Begriffsschrift, Frege makes the following remarks about his work (the following translations are by J. van Heijenoort [12], p. 6-7).

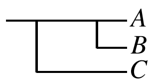
> (p. X) Its first purpose, therefore, is to provide us with the most reliable test of the validity of a chain of inferences and to point out every presupposition that tries to sneak in unnoticed, so that its origin can be investigated. That is why I decided to forgo expressing anything that is without significance for the inferential sequence. In § 3 I called what alone mattered to me the *conceptual content* [begrifflichen Inhalt].
>
> (p.XI) I believe that I can best make the relation of my ideography to ordinary language [Sprache des Lebens] clear if I compare it to that which the microscope has to the eye. Because of the range of its possible uses and the versatility with which it can adapt to the most diverse circumstances, the eye is far superior to the microscope. Considered as an optical instrument, to be sure, it exhibits many imperfections, which ordinarily remain unnoticed only on account of its intimate connection with our mental life. But, as soon as scientific goals demand great sharpness of resolution, the eye proves to be insufficient. The microscope, on the other hand, is perfectly suited to precisely such goals, but that is just why it is useless for all others.
>
> (p.XII) If it is one of the tasks of philosophy to break the domination of the word over the human spirit by laying bare the misconceptions that through the use of language often almost unavoidably arise concerning the relations between concepts and by freeing thought from that with which only the means of expression of ordinary language, constituted as they are, saddle it, then my ideography, further developed for these purposes, can become a useful tool for the philosopher. To be sure, it too will fail to reproduce ideas in a pure form, and this is probably inevitable when ideas are represented by concrete means; but, on the one hand, we can restrict the discrepancies to those that are unavoidable and harmless, and, on the other, the fact that they are of a completely different kind from those peculiar to ordinary language already affords protection against the specific influence that a particular means of expression might exercise. [J. van Heijenoort [12], p. 6-7]

The notation that Frege introduces in his Begriffsschrift has not survived. It presents difficulties in printing and takes up a large amount of space. But, as Frege himself says, 'the comfort of the typesetter is certainly not the summum bonum, and the notation undoubtedly allows one to perceive the structure of a formula at a glance and to perform substitutions with ease.'

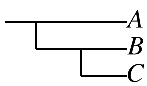
In § 5 of his Begriffsschrift Frege introduces the notation



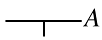
for our $B \to A$. Our $C \to (B \to A)$ is represented by Frege as:



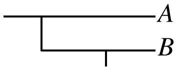
while Frege represents our $(C \to B) \to A$ by:



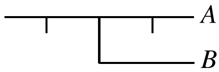
In section 7 of his Begriffsschrift Frege represents our $\neg A$ by:



Frege presents the propositional calculus in a version that uses the conditional and negation as primitive connectives. Frege renders our $A \vee B$ by $\neg B \to A$, i.e.,



And Frege renders our $A \wedge B$ by $\neg(B \to \neg A)$, i.e.,



The distinction between 'and' and 'but' is of the kind that is not expressed in the present ideography. [G. Frege, *Begriffsschrift*, § 7.]

Conversational implicature P. Grice in the 1967 William James Lectures (published in 1989 in [10]) works out a theory in *pragmatics* which he calls the theory of *conversational implicature*. Generally speaking, in conversation we usually obey or try to obey rules something like the following:

QUANTITY: Be informative
QUALITY: Tell the truth
RELATION: Be relevant
MODE: Avoid obscurity, prolixity, etc.

If the fact that A has been said, plus the assumption that the speaker is observing the above rules, plus other reasonable assumptions about the speaker's purposes and intentions in the context, logically entails that B, then we can say *A conversationally implicates B*.

It is possible for A to conversationally implicate many things which are in no way part of the *meaning* of A. For example, if X says 'I'm out of gas' and Y says 'there's a gas station around the corner', Y's remark conversationally implicates that the station in question is open, since the information that the station is there would be *irrelevant* to X's predicament otherwise. If X says 'Your hat is either upstairs in the back bedroom or down in the hall closet', this remark conversationally implicates 'I don't know which', since if X did know which, this remark would not be the most *informative* one he could provide.

Grice shows how philosophers have sometimes mistaken conversational implicatures for elements of meaning. For instance, Strawson sometimes claims not-knowing-which must be part of the *meaning* of 'or' (and therefore the traditional treatment of disjunction in logic is misleading or false). Grice claims this is mistaking the conversational implicature cited above for an aspect of meaning.

Sometimes it is possible to *cancel* a conversational implicature by adding something to one's remark. For example, in the gas station case, 'I'm not sure whether it's open' and in the hat case, 'I know, but I'm not saying which' (one might say this if locating the hat was part of some sort of parlor game). The possibility of cancellation shows that the conversational implicatures definitely are not part of the *meaning* of the utterance.

Conditionals In the examples below the conditional in (1) is in the *indicative* mood, while the conditional in (2) is a *subjunctive* one.
(1) If Oswald did not kill Kennedy, someone else did.
(2) If Oswald had not killed Kennedy, someone else would have.
(These examples are from E. Adams, *Subjunctive and Indicative Conditionals*, Foundations of Language 6: 89-94, 1970.)

(1) is true: someone killed Kennedy; but (2) is probably false. Therefore, different analyses are needed for indicative and for subjunctive conditionals.

A *counterfactual* conditional is an expression of the form 'if A were the case, then B would be the case', where A is supposed to be false. Not all subjunctive conditionals are counterfactual. Consider the argument, 'The murderer used an ice pick. But if the butler had done it, he wouldn't have used an ice pick. So the murderer must have been someone else.'. If this subjunctive conditional were a counterfactual, then the speaker would be presupposing that the conclusion of his argument is true. (This example is from R.C. Stalnaker, *Indicative Conditionals*, in W.L. Harper, e.a., *IFS*.)

In Chapter 6 we shall discuss counterfactuals and subjunctive conditionals in general. In this section we will restrict our attention from now on to indicative conditionals.

In Section 2.4 we have considered the so-called *paradoxes of material implication*: the following two inferences for material implication $\rightarrow$ are valid, whereas the corresponding English versions seem invalid.

$$\frac{\neg A}{A \rightarrow B} \qquad \frac{\text{There is no oil in my coffee}}{\text{If there is oil in my coffee, then I like it}}$$

$$\frac{B}{A \rightarrow B} \qquad \frac{\text{I'll ski tomorrow}}{\text{If I break my leg today, then I'll ski tomorrow}}$$

So, the truth-functional reading of 'if ..., then ...', in which $A \rightarrow B$ is equivalent to $\neg A \vee B$, seems to conflict with judgments we ordinarily make. The paradoxical character of these inferences disappears if one realizes that:
1. the material implication $A \rightarrow B$ has the same truth-table as $\neg A \vee B$;
2. speaking the truth is only one of the conversation rules one is expected to obey in daily discourse; one is also expected to be as relevant and informative as possible.
Now, if one has at one's disposal the information $\neg A$ (or B, respectively) and at the same time provides the information $A \rightarrow B$, i.e., $\neg A \vee B$, then one is speaking the truth, but a truth calculated to mislead, since the premiss $\neg A$ (or B, respectively) is so much simpler and more informative than the conclusion $A \rightarrow B$. If one knows the premiss $\neg A$ (or B, respectively), the conversation rules force us to assert this premiss instead of $A \rightarrow B$. Quoting R. Jeffrey:

> Thus defenders of the truth-functional reading of everyday conditionals point out that the disjunction $\neg A \vee B$ shares with the conditional 'if A, then B' the feature that normally it is not to be asserted by someone who is in a position to deny A or to assert B. ...
>
> Normally, then, conditionals will be asserted only by speakers who think the antecedent false or the consequent true, but do not know which. Such speakers will think they know of some connection between the components, by virtue of which they are sure (enough for the purposes at hand) that the first is false or the second is true. [R. Jeffrey [13], pp. 77-78]

Summarizing in a slogan:

indicative conditional = material implication + conversation rules.

So H.P. Grice uses principles of conversation to explain facts about the use of conditionals that seem to conflict with the truth-functional analysis of the ordinary indicative conditional. In his paper 'Indicative Conditionals' (in W.L. Harper, e.a. (eds.), *IFS*), R.C. Stalnaker follows another strategy, rejecting the material conditional analysis. And in his book 'Causal Necessity', Brian Skyrms claims that the indicative conditional cannot be construed as the material implication '$\rightarrow$' plus conversational implicature. The dispute between advocates of the truth-functional account of conditionals and the advocates of other, more complex but seemingly more adequate accounts is as old as logic itself.

Frege, Russell, Hilbert In his *Begriffsschrift* (page 2) of 1879 Gottlob Frege distinguishes the notations $–A$ for 'the proposition that A' and $\vdash A$ for 'it is a fact that A'. Frege calls A in $–A$ and in $\vdash A$ 'der Inhalt' (the *content*) and '$\vdash A$' 'ein Urteil' (a *judgment*). In Chapter II of his book Frege gives the first axiomatic formulation of classical propositional (and predicate) logic, namely, the following system $\mathscr{P}_F$, presented below in our own notation.

$A \rightarrow (B \rightarrow A)$	(Begriffsschrift, p. 26, form. 1)
$(C \rightarrow (B \rightarrow A)) \rightarrow ((C \rightarrow B) \rightarrow (C \rightarrow A))$	(Begriffsschrift, p. 26, form. 2)
$(D \rightarrow (B \rightarrow A)) \rightarrow (B \rightarrow (D \rightarrow A))$	(Begriffsschrift, p. 35, form. 8)
$(B \rightarrow A) \rightarrow (\neg A \rightarrow \neg B)$	(Begriffsschrift, p. 43, form. 27)
$\neg\neg A \rightarrow A$	(Begriffsschrift, p. 44, form. 31)
$A \rightarrow \neg\neg A$	(Begriffsschrift, p. 47, form. 41)

together with Modus Ponens.

It is probably correct to say that Frege's work only became well-known through Russell. The following formulation $\mathscr{P}_R$ of classical propositional logic was used by Whitehead and Russell in *Principia Mathematica* in 1910 (see part I, page 13).

$$A \vee A \rightarrow A$$
$$B \rightarrow A \vee B$$
$$A \vee B \rightarrow B \vee A$$
$$A \vee (B \vee C) \rightarrow B \vee (A \vee C)$$
$$(B \rightarrow C) \rightarrow (A \vee B \rightarrow A \vee C)$$

together with Modus Ponens.

The following formulation $\mathscr{P}$ of propositional logic has implication and negation as primitive connectives and Modus Ponens as its only rule:

$$A \rightarrow (B \rightarrow A)$$
$$(A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C))$$
$$(\neg A \rightarrow \neg B) \rightarrow (B \rightarrow A)$$

Defining $A \wedge B := \neg(A \rightarrow \neg B)$ and $A \vee B := (A \rightarrow B) \rightarrow B$, the axioms for $\wedge$ and $\vee$ in Section 2.6 become formulas containing no connectives other than $\rightarrow$ and $\neg$ and are deducible (using *MP*) from the three axiom schemes given above. So, by expressing $\wedge$ and $\vee$ in terms of $\rightarrow$ and $\neg$, formulations such as $\mathscr{P}$ are obtained, in which the number of axioms is small.

In their *Grundlagen der Mathematik* (1934) D. Hilbert (1862-1943) and P. Bernays (1888-1977) presented the following axiom system $\mathscr{P}_H$ for the classical propositional calculus. This system contains axioms for each of the connectives $\rightarrow$, $\wedge$, $\vee$ and $\neg$.

$$\left.\begin{array}{l} A \to (B \to A) \\ (A \to (B \to C)) \to ((A \to B) \to (A \to C)) \end{array}\right\} \to$$

$$\left.\begin{array}{l} A \wedge B \to A \\ A \wedge B \to B \\ A \to (B \to A \wedge B) \end{array}\right\} \wedge$$

$$\left.\begin{array}{l} A \to A \vee B \\ B \to A \vee B \\ (A \to C) \to ((B \to C) \to (A \vee B \to C)) \end{array}\right\} \vee$$

$$(\neg A \to \neg B) \to (B \to A) \qquad \neg$$

Formulations of *intuitionistic propositional logic* can be obtained by replacing the negation axiom of $\mathscr{P}_H$ by suitable different axioms, for instance, by $(A \to \neg A) \to \neg A$ and $\neg A \to (A \to B)$; see Chapter 8.

For more historical details the reader is referred to section 29 of A. Church [4]. *Introduction to Mathematical Logic.*

Scientific Explanation, Inductive Logic Some, but not all, scientific explanations are deductive arguments the premisses of which consist of general laws and particular facts. A trivial example is the following explanation.

If someone drops his pencil, it falls to the ground. (L_1)
I drop my pencil. (C_1)
Therefore, my pencil falls to the ground. (E)

L_1 is a *general law*, i.e., a universal statement expressing that each time some condition P is satisfied, then without exception some condition Q will occur. C_1 is a *particular fact*. And E is the *explanandum*, the statement which has to be explained.

Explanations of this kind are called *deductive-nomological explanations*. (The Greek word 'nomos' means 'law'.) Their general form is

$$\left.\begin{array}{ll} L_1, L_2, \ldots, L_r & \text{(universal laws)} \\ C_1, C_2, \ldots, C_k & \text{(particular facts)} \end{array}\right\} \text{Explanans}$$
$$\overline{\qquad E \qquad} \qquad \text{Explanandum}$$

In a deductive-nomological explanation the explanandum follows logically or deductively from the explanans.

Probabilistic explanations are different in that i) the laws are in terms of relative frequences, and ii) the explanandum does not logically follow from the explanans, but can only be expected with a certain degree of probability, called *inductive* or *logical probability*. The following is an example of a probabilistic explanation.

Example 2.29. The statistical probability of catching the measles, when exposed to them, is $\frac{3}{4}$. The statistical probability of catching pneumonia, when exposed to it, is $\frac{1}{7}$. Jim was exposed to the measles and to pneumonia. Therefore, the inductive or logical probability that Jim catches both the measles and pneumonia is $\frac{3}{4} \times \frac{1}{7} = \frac{3}{28}$.

The main question in *inductive logic* is how to determine the inductive probability for the explanandum, given the statistical probabilities in the explanans. This problem is in part still unsettled. Note that inductive or logical probability is a relation between statements, while statistical probability is a relation between (kinds of) events.

References for further reading: 1. Hempel, R., *Philosophy of Natural Science*; 2. Carnap, R., *Logical foundations of probability*; 3. Carnap, R. and Jeffrey, R., *Studies in inductive logic and probability*; 4. Jeffrey, R., *The logic of decision*; 5. Swinburne, R., *An introduction to confirmation theory*.

Syntax – Semantics The *syntax* of a language is concerned only with the *form* of the expressions, while the *semantics* is concerned with their *meaning*.

So, the rules according to which the well-formed expressions of a language are formed and the rules belonging to a logical proof system, such as Modus Ponens, belong to the syntax of the language in question. These rules can be manipulated mechanically; a machine can be instructed to apply the rule Modus Ponens and to write down a B once it sees both A and $A \to B$, while the machine does not know the meanings of A, B and $\to$. The notions of (logical) proof and deduction, as well as the notions of (logical) provability and deducibility, clearly belong to the syntax: they are only concerned with the form of the formulas involved.

On the other hand, truth tables belong to the semantics, because they say how the truth value (meaning) of a composite proposition is related to the truth values (meanings) of the components from which it is built. The notions of validity and valid consequence also belong to the semantics: they are concerned with the meaning of the formulas in question.

Leibniz (1646-1716) We will here pay attention to only a few aspects of Leibniz. For more information the reader is referred to Kneale [16] and to Mates [20], Chapter 12. What follows in this subsection is based on these works.

One of Leibniz' ideals was to develop a *lingua philosophica* or *characteristica universalis*, an artificial language that in its structure would mirror the structure of thought and that would not be affected with ambiguity and vagueness like ordinary language. His idea was that in such a language the linguistic expressions would be pictures, at it were, of the thoughts they represent, such that signs of complex thoughts are always built up in a unique way out of the signs for their composing parts. Leibniz believed that such a language would greatly facilitate thinking and communication and that it would permit the development of mechanical rules for deciding all questions of consistency or consequence. The language, when it is perfected, should be such that 'men of good will desiring to settle a controversy on any subject whatsoever will take their pens in their hands and say *Calculemus* (let us calculate)'. If we restrict ourselves to propositional logic, Leibniz' ideal has been realized: classical propositional logic is decidable; see Section 2.9. However, A. Church and A. Turing proved in 1936 that extending the propositional language with the quantifiers 'for all' ($\forall$) and 'for some' ($\exists$), the resulting predicate logic is undecidable, i.e., there is no mechanical method to test logical consequence (in predicate logic), let alone philosophical truth.

Leibniz also developed a theory of identity, basing it on *Leibniz' Law*: eadem sunt quorum unum potest substitui alteri salva veritate – those things are the same if one may be substituted for the other with preservation of truth. Leibniz' Law is also called the *substitutivity of identity* or the *principle of extensionality* and it is frequently formulated as follows.

$$a = b \rightarrow (\ldots a \ldots \rightleftarrows \ldots b \ldots)$$

where $\ldots a \ldots$ is a context containing occurrences of the name a, and $\ldots b \ldots$ is the same context in which one or more occurrences of a has been replaced by b; if a is b, then what holds for a holds for b and vice versa. In the propositional calculus we have a similar principle of the *substitutivity of material equivalents*:

$$(A \rightleftarrows B) \rightarrow (\ldots A \ldots \rightleftarrows \ldots B \ldots).$$

Leibniz made a distinction between *truths of reason* and *truths of fact*. The truths of reason are those which could not possibly be false, i.e., – in modern terminology – which are *necessarily true* . Examples of such truths are: no bachelor is married, $2+2=4$, living creatures cannot survive fire, and so on. Truths of fact are called contingent truths nowadays; for example, unicorns do not exist, Amsterdam is the capital of the Netherlands, and so on. Leibniz spoke of the truths of reason as *true in all possible worlds*. He imagined that there are many *possible worlds* and that our actual world is one of them. '$2+2=4$' is true not only in this world, but also in any other world. 'Amsterdam is the capital of the Netherlands' is true in this world, but we can think of another world in which this proposition is false. In 1963, S. Kripke extended the notion of possible world with an *accessibility relation* between possible worlds, which enabled him to give adequate semantics for the different modal logics, as we will see in Chapter 6. The idea is that some worlds are accessible from the given world, and some are not. For instance, one could postulate (and one usually does) that worlds with different mathematical laws are not accessible from the present world.

2.11 Solutions

Solution 2.1. i) $P_1 \wedge P_2 \rightarrow \neg P_3$; ii) $P_1 \wedge (P_2 \rightarrow \neg P_3)$; iii) $P_1 \vee (P_2 \rightarrow P_3)$; iv) $(P_2 \vee P_1) \rightarrow P_3$; v) $P_1 \rightarrow (P_2 \rightarrow \neg P_3)$

Solution 2.2. i) If it is the case that if John works hard then he goes to school, then John is not wise. ii) John does not work hard or John is wise. iii) It is not the case that John works hard or that John is wise; in other words, John does not work hard and John is not wise. iv) John does not go to school and John is wise. v) It is not the case that both John goes to school and John is wise; in other words, John does not go to school or John is not wise.

Solution 2.3. 1. P_1 or $\forall x[P(x)]$; 2. P_2 or $\forall x[\neg P(x)]$; 3. $\neg P_1$ or $\neg\forall x[P(x)]$.

Solution 2.4. Only the expressions P_1, $\neg P_8$, $P_1 \wedge \neg P_8$, and $(P_1 \wedge P_2) \rightarrow \neg P_3$ are formulas of propositional logic. All other expressions contain symbols which do not occur in the alphabet of propositional logic.

Solution 2.5. Let Φ be the property defined by $\Phi(n) := 1+2+\ldots n = \frac{1}{2}n(n+1)$.
1. 0 has the property Φ, since $0 = \frac{1}{2}0(0+1)$.
2. Suppose n has the property Φ, i.e., $1+2+\ldots n = \frac{1}{2}n(n+1)$ (induction hypothesis). Then we have to show that $n+1$ also has the property Φ, i.e., $1+2+\ldots n+(n+1) = \frac{1}{2}(n+1)((n+1)+1)$.
Proof: According to the induction hypothesis, $1+2+\ldots n+(n+1) = \frac{1}{2}n(n+1)+(n+1) = (\frac{1}{2}n+1)(n+1) = \frac{1}{2}(n+1)(n+2)$.

Solution 2.6. Atomic formulas have no or zero parentheses, so as many left parentheses as right parentheses.
Assume that A and B have as many left parentheses as right parentheses (induction hypothesis). Then evidently the formulas $(A \rightleftarrows B)$, $(A \rightarrow B)$, $(A \wedge B)$, $(A \vee B)$ and $(\neg A)$ also have as many left parentheses as right parentheses.

Solution 2.7. We restrict ourselves to showing that $\neg(A \wedge B)$ has the same truth table as $\neg A \vee \neg B$. Although the formulas A and B may have been composed of many atomic formulas $P_1, \ldots, P_n$ and hence their truth tables may consist of many lines, 2^n, in the end there are at most 4 possible different combinations of 1 (true) and 0 (false) for A and B. Hence, it suffices to restrict ourselves to these maximally 4 possible different combinations:

A	B	$A \wedge B$	$\neg(A \wedge B)$	$\neg A$	$\neg B$	$\neg A \vee \neg B$
1	1	1	0	0	0	0
1	0	0	1	0	1	1
0	1	0	1	1	0	1
0	0	0	1	1	1	1

Solution 2.8. Below are the truth tables for the formulas from exercise 2.1 and 2.2.

			2.1					2.2				
P_1	P_2	P_3	i	ii	iii	iv	v	i	ii	iii	iv	v
1	1	1	0	0	1	1	0	0	1	0	0	0
1	1	0	1	1	1	0	1	1	0	0	0	1
1	0	1	1	1	1	1	1	1	1	0	1	1
1	0	0	1	1	1	0	1	1	0	0	0	1
0	1	1	1	0	1	1	1	0	1	0	0	0
0	1	0	1	0	0	0	1	1	1	1	0	1
0	0	1	1	0	1	1	1	0	1	0	1	1
0	0	0	1	0	1	1	1	1	1	1	0	1

Solution 2.9. $A \vee \neg A$ has the value 1 and $A \wedge \neg A$ has the value 0 in all lines of the truth table. Hence, in each line of the truth table
a) $(A \vee \neg A) \rightarrow B$ is 0 iff B is 0,
b) $(A \vee \neg A) \wedge B$ is 0 iff B is 0, and
c) $(A \wedge \neg A) \vee B$ is 0 iff B is 0.
Therefore $(A \vee \neg A) \rightarrow B$, $(A \vee \neg A) \wedge B$ and $(A \wedge \neg A) \vee B$ have the same truth table as B.

Solution 2.10.

A	B	$A \to B$	$(A \to B) \to B$	$B \to A$	$(B \to A) \to A$	$A \vee B$
1	1	$1 \to 1 = 1$	$1 \to 1 = 1$	1	$1 \to 1 = 1$	1
1	0	$1 \to 0 = 0$	$0 \to 0 = 1$	1	$1 \to 1 = 1$	1
0	1	$0 \to 1 = 1$	$1 \to 1 = 1$	0	$0 \to 0 = 1$	1
0	0	$0 \to 0 = 1$	$1 \to 0 = 0$	1	$1 \to 0 = 0$	0

Alternatively, one might argue as follows: $(A \to B) \to B$ is 0 iff ($A \to B$ is 1 and B is 0) iff (A is 0 and B is 0) iff $A \vee B$ is 0. Similarly for $(B \to A) \to A$.

Solution 2.11. We restrict ourselves to c) $(P \to Q) \to (\neg Q \to \neg P)$. Suppose in some line of its truth table this formula has the value 0. Then in that line $P \to Q$ is 1 and $\neg Q \to \neg P$ is 0. Hence, $P \to Q$ is 1, $\neg Q$ is 1 and $\neg P$ is 0 in that same line. So, $P \to Q$ is 1, Q is 0 and P is 1 in that same line. Then, either P is 0, Q is 0 and P is 1 in that line, or Q is 1, Q is 0 and P is 1 in that line. Both are impossible, so the original formula cannot have the value 0 in some line of its truth table.

Solution 2.12. a) In order that $P \vee Q \to P \wedge Q$ is 0 in some line, $P \vee Q$ must be 1 and $P \wedge Q$ 0 in that same line. So, at least one of P, Q must be 0. By taking the value of the other formula 1, one achieves that $P \vee Q$ is 1, while $P \wedge Q$ is 0:

P	Q	$P \vee Q$	$P \wedge Q$
1	0	1	0
0	1	1	0

b) is treated similarly.

Solution 2.13. 1 B, 2 A, 3 B, 4 C, 5 B, 6 C, 7 C, 8 C, 9 C, 10 C.

Solution 2.14. Each formula A built by means of connectives from only one atomic formula P must have one of the following four truth tables.

P	A			
1	1	1	0	0
0	1	0	1	0

These four truth tables are the tables of $P \to P$, P, $\neg P$ and $P \wedge \neg P$, respectively.

Solution 2.15. Straightforward

Solution 2.16. * Let G be a group. If G can be ordered, then clearly every subgroup of G, generated by finitely many elements of G, can be ordered. Conversely, suppose every such subgroup of G can be ordered. $(*)$

Now, consider the propositional language built from atomic formulas $P_{a,b}$, where a, b are elements of G. Let Γ be the following set of formulas in this language.

$P_{a,a}$ for every element a in G.

$P_{a,b} \vee P_{b,a}$ for all a, b in G.

$P_{a,b} \to \neg P_{b,a}$ for all a, b in G with $a \neq b$.

$P_{a,b} \wedge P_{b,c} \to P_{a,c}$ for all a, b, c in G.

$P_{a,b} \to P_{ac,bc} \wedge P_{ca,cb}$ for all a, b, c in G.

Proposition 1: Every finite subset of Γ has a model.

Proof: Let Γ' be a finite subset of Γ. In Γ' there occur only finitely many elements of G. Let G' be the subgroup of G, generated by these finitely many elements. By the hypothesis $(*)$, G' can be ordered by some relation $\leq$. Now, let $u(P_{a,b}) = 1$ if $a \leq b$, and $u(P_{a,b}) = 0$ if $a > b$. Then u is a model of Γ'.
By the compactness theorem it follows from Proposition 1 that Γ has a model, say v. Now, let $a \leq b := v(P_{a,b}) = 1$. Since v is a model of Γ, $\leq$ is an ordering of G.

Solution 2.17. * If a graph on V is k-chromatic, then clearly every finite subgraph of it is k-chromatic. Conversely, suppose R is a graph on V such that every finite sub-graph of R is k-chromatic. $(*)$
Now, consider the propositional language built from atomic formulas $P_{i,x}$, where $i \in \{1,\ldots,k\}$ and $x \in V$. And let Γ be the following set of formulas.
$P_{i,x} \to \neg P_{j,x}$ for all $i, j \leq k$ with $i \neq j$ and all $x \in V$.
$P_{1,x} \vee \ldots \vee P_{k,x}$ for all $x \in V$.
$P_{i,x} \to \neg P_{i,y}$ for all $i \leq k$ and all $x,\ y \in V$ such that xRy.
Proposition 1: Every finite subset of Γ has a model.
Proof: Let Γ' be a finite subset of Γ. In Γ' there occur only finitely many elements of V. Let R' be the sub-graph of R obtained by restricting R to the set V' of these finitely many elements. By hypothesis $(*)$, R' is k-chromatic, i.e., there is a partition of V' into k disjoint sets $W_1,\ldots,W_k$, such that two elements of V' connected by R' do not belong to the same W_i. Now, let $u(P_{i,x}) = 1$ if $x \in W_i$, and $u(P_{i,x}) = 0$ if $x \notin W_i$. Then u is a model of Γ'.

By the compactness theorem it follows from proposition 1 that Γ has a model, say v. Now, let $V_i := \{x \in V \mid v(P_{i,x}) = 1\}$ for $i = 1,\ldots,k$. Then $V_1,\ldots,V_k$ is a partition of V such that two elements of V, connected by R, do not belong to the same V_i. In other words, R is k-chromatic.

Solution 2.18. * Let B and G be sets. $R \subseteq B \times G$, such that (i) for all $x \in B$, $R_{\{x\}}$ is finite, and (ii) for every finite subset $B' \subseteq B$, $R_{B'}$ has at least as many elements as B'. Consider a propositional language with as atomic formulas all expressions $H_{x,y}$ with $x \in B$ and $y \in G$. Let Γ contain the following formulas:
$H_{x,y_1} \vee \ldots \vee H_{x,y_n}$ for any $x \in B$, where $R_{\{x\}} = \{y_1,\ldots,y_n\}$.
$\neg(H_{x,y_1} \wedge H_{x,y_2})$ for any $x \in B$, $y_1, y_2 \in G$ with $y_1 \neq y_2$.
$\neg(H_{x_1,y} \wedge H_{x_2,y})$ for any $x_1, x_2 \in B$, $y \in G$ with $x_1 \neq x_2$.
If u is a model of Γ, then $f : B \to G$, defined by $f(x) = y$ if $u(H_{x,y}) = 1$, is an injection from B to G.

In order to show that Γ has a model, by the compactness theorem it suffices to show that each finite subset Γ' of Γ has a model. So, let Γ' be a finite subset of Γ. Let $B' := \{x \in B \mid H_{x,y}$ occurs in Γ' for some $y \in G\}$, and $G' := \{y \in G \mid H_{x,y}$ occurs in Γ' for some $x \in B\}$. Since B' and G' are finite, there is an injection $f' : B' \to G'$, such that if $f'(x) = y$, then $R(x,y)$. Define u' as follows: $u'(H_{x,y}) = 1$ iff $f'(x) = y$. Then u' is a model of Γ'.

Solution 2.19.

P_1	P_2	P_3	$P_1 \to P_2$	$\neg P_2 \vee P_3$	$P_1 \to P_3$	$P_3 \to P_1$
1	1	1	**1**	**1**	**1**	**1**
1	1	0	1	0	0	1
1	0	1	0	1	1	1
1	0	0	0	1	0	1
0	1	1	**1**	**1**	**1**	**0**
0	1	0	1	0	1	1
0	0	1	**1**	**1**	**1**	**0**
0	0	0	**1**	**1**	**1**	**1**

Let P_1 stand for: the government raises taxes for its citizens; P_2 for: the unemployment grows; and P_3 for: the income of the state decreases. Then the argument has the following structure: $P_1 \to P_2, \neg P_2 \vee P_3 \models P_1 \to P_3$. Notice that $\neg P_2 \vee P_3$ has the same truth table as $P_2 \to P_3$. One easily checks that in each line of the truth table starting with P_1, P_2, P_3 in which both premisses are 1, also the conclusion is 1.

There are four lines in which all premisses are true: line 1, 5, 7 and 8. In each of these lines the conclusion $P_1 \to P_3$ is 1 too. Therefore, $P_1 \to P_2, \neg P_2 \vee P_3 \models P_1 \to P_3$.

Solution 2.20. Let P_1 stand for: Europe may form a monetary union; P_2 for: Europe is a political union; and P_3 for: all European countries are member of the union. Then the argument has the following structure: $P_1 \to P_2, \neg P_2 \vee P_3 \models P_3 \to P_1$, which is false, because there is at least one line in the truth table in which all premisses are 1, while the putative conclusion $P_3 \to P_1$ is 0; see lines 5 and 7 in the table of solution 2.19. Therefore, $P_1 \to P_2, \neg P_2 \vee P_3 \not\models P_3 \to P_1$.

Solution 2.21. c) There is no line in the truth table in which both A and $\neg A$ are 1, so there is no line in the truth table in which both A and $\neg A$ are 1 and B is 0, i.e., $A, \neg A \models B$.

Solution 2.22. Let W stand for: John wins the lottery; J for: John makes a journey; and S for: John succeeds for logic. Then the structure of the argument is the following one: $\neg W \vee J,\ \neg J \to \neg S,\ W \vee S \models J$. Notice that the first premiss has the same truth table as $W \to J$ and that the second premiss has the same truth table as $S \to J$. Hence, the structure of the argument is equivalent to $W \to J,\ S \to J,\ W \vee S \models J$, which clearly is valid. Checking the truth table will confirm this.

Solution 2.23. Let T stand for: Turkey joins the EU; L for: the EU becomes larger; and S for: the EU becomes stronger. Then the argument has the following structure: $T \to L,\ \neg(S \wedge \neg L) \models \neg T \vee S$. Notice that $\neg(S \wedge \neg L)$ has the same truth table as $S \to L$ and that the conclusion $\neg T \vee S$ has the same truth table as $T \to S$. Hence the structure of the argument is equivalent to $T \to L,\ S \to L \models T \to S$, which clearly does not hold: if T and L are 1 and S is 0, then the premisses are both 1, while the conclusion is 0. Making a truth table will confirm this.

Solution 2.24. 1) Assume $\models A \rightleftarrows (A \to B)$. To show: $\models A$ and $\models B$. So, suppose A were 0 in some line of its truth table. Then $A \rightleftarrows (A \to B)$ would be $0 \rightleftarrows (0 \to 0/1) = (0 \rightleftarrows 1) = 0$ in that line, contradicting the assumption. Therefore, $\models A$. In a similar

way $\models B$ can be shown.
2) Assume $A \models \neg A$. To show: $\models \neg A$. So, suppose $\neg A$ were 0 in some line of its truth table, i.e. A were 1 in that line. Then, by assumption, also $\neg A$ would be 1 in that same line. Contradiction. Therefore, $\models \neg A$.
3) Assume $A \to B \models A$. To show: $\models A$. So, suppose A were 0 in some line of its truth table. Then $A \to B$ would be 1 in that same line and hence, by assumption, A would be 1 in that same line. Contradiction. Therefore, $\models A$.

Solution 2.25. a) Counterexample: let $A = P$ (atomic) and $B = Q$ (atomic). Then not $\models P \to Q$, but not $\models P$ and not $\models \neg Q$.
b) Proof: $\neg(A \to B)$ has the same truth table as $A \wedge \neg B$. So, if $\models \neg(A \to B)$, then $\models A \wedge \neg B$. Hence, by Theorem 2.14, $\models A$ and $\models \neg B$.
c) Counterexample: let $A = P$ (atomic) and $B = Q$ (atomic). Then not $\models P \wedge Q$, but not $\models \neg P$ and not $\models \neg Q$.
d) Counterexample: let $A = P$ (atomic) and $B = \neg P$. Then $\models \neg(P \wedge \neg P)$, but not $\models \neg P$ and not $\models \neg\neg P$. Notice that $A = P$ and $B = Q$ with P, Q atomic, is not a counterexample, because $\models \neg(P \wedge Q)$ does not hold.
e) Counterexample: $A = P$ (atomic) and $B = Q$ (atomic). Then not $\models P \vee Q$, but not $\models \neg P$ and not $\models \neg Q$.
f) Proof: $\neg(A \vee B)$ has the same truth table as $\neg A \wedge \neg B$. So, if $\models \neg(A \vee B)$, then $\models \neg A \wedge \neg B$. Hence, by Theorem 2.14, $\models \neg A$ and $\models \neg B$.

Solution 2.26. (a1) and (a2) For $i = 1, \ldots, n$, $A_1, \ldots, A_i, \ldots, A_n \models A_i$, since for every line in the truth table, if all of $A_1, \ldots, A_i, \ldots, A_n$ are 1, then also A_i is 1.
(b1) Assume $A_1, A_2, A_3 \models B_1$ and $A_1, A_2, A_3 \models B_2$ and $B_1, B_2 \models C$, i.e., for every line in the truth table, if all of A_1, A_2, A_3 are 1, then also B_1 is 1 and B_2 is 1; and for every line in the truth table, if all of B_1, B_2 are 1, then also C is 1. Therefore, for every line in the truth table, if all of A_1, A_2, A_3 are 1, then also C is 1, i.e., $A_1, A_2, A_3 \models C$.
(b2) Similarly.

Solution 2.27. 1) Assume $A \models B$ and $A \models \neg B$ and suppose that in some line of the truth table $\neg A$ is 0, i.e., A is 1. Then, because of $A \models B$, B is 1 in that line and, because of $A \models \neg B$, $\neg B$ is 1 (and hence B is 0) in that line of the truth table. Contradiction. So, there is no line in which A is 1. Therefore $\models \neg A$.
2) Assume $A \models C$ and $B \models C$ and, in order to show that $A \vee B \models C$, suppose $A \vee B$ is 1 in some line of the truth table. Then A is 1 or B is 1 in that line. In the first case it follows from $A \models C$ and in the second case it follows from $B \models C$ that C is 1 in that line.

Solution 2.28. (a) Right. There is no line in the truth table in which $A \to B \vee C$ is 1 and $(A \to B) \vee (A \to C)$ is 0.
(b) Wrong. Counterexample: for P, Q atomic, $\models (P \to Q) \vee (P \to \neg Q)$, but not $\models P \to Q$ and not $\models P \to \neg Q$. (See also Theorem 2.13 (b))
(c) Assume $A \models B$ (1). To show: $B \to C \models A \to C$. So, suppose $B \to C$ is 1 in some line of the truth table (2). Then we have to show that also $A \to C$ is 1 in that line. So, suppose A is 1 in that same line (3). Then, because of (1), B is 1 in that line and hence, because of (2), C is 1 in that line, which had to be proved.

Solution 2.29. Assume $T \wedge A \wedge B \models P$. To show: $\neg P \models \neg T \vee \neg A \vee \neg B$. So, suppose $\neg P$ is 1 in some line of the truth table. Then P is 0 in that line and hence, by assumption, $T \wedge A \wedge B$ is 0 in that line. Then $\neg(T \wedge A \wedge B)$ is 1 and hence $\neg T \vee \neg A \vee \neg B$ is 1 in the given line. Therefore, $\neg P \models \neg T \vee \neg A \vee \neg B$.

Solution 2.30. Proof of a): Assume $A \models B$. To show: $\neg B \models \neg A$. So, suppose $\neg B$ is 1 in an arbitrary line of the truth table. Then B is 0 in that line and hence, by assumption, A is 0 in that line. Therefore $\neg A$ is 1 in that line, which had to be shown.
Proof of b): Assume $A \models B$ and $A, B \models C$. To show $A \models C$. So, suppose A is 1 in an arbitrary line of the truth table. Then, because of $A \models B$, A and B are 1 in that line and hence, by $A, B \models C$, C is 1 in that line, which had to be shown.
Proof of c): Assume $A \vee B \models A \wedge B$. And suppose A and B have different values in some line of the truth table ($1-0$ or $0-1$ respectively). Then $A \vee B$ is 1 in that line, while $A \wedge B$ is 0 in that line, contradicting $A \vee B \models A \wedge B$. Therefore A and B have the same truth table.

An alternative proof: Suppose $A \vee B \models A \wedge B$. This means that the formulas A and B are such that in the standard truth table for $A \vee B$ and for $A \wedge B$ line 2 (A is 1, B is 0) and line 3 (A is 0 and B is 1) do not occur. So, only line 1 (A is 1 and B is 1) and line 4 (A is 0 and B is 0) may occur. Hence, A and B have the same truth table.

Solution 2.31.

			Brown's testimony	Jones' testimony	Smith's testimony
B	J	S	$\neg J \wedge S$	$\neg B \to \neg S$	$S \wedge (\neg B \vee \neg J)$
1	1	1	0	1	0
1	1	0	0	1	0
1	0	1	1	1	1
1	0	0	0	1	0
0	1	1	0	0	1
0	**1**	**0**	**0**	**1**	**0**
0	0	1	1	0	1
0	0	0	0	1	0

a) Yes, for the three testimonies are all true in the third line of the truth table.
b) $\neg J \wedge S \models S \wedge (\neg B \vee \neg J)$, i.e., Smith's testimony follows from that of Brown.
c) The assumption that everybody is innocent means in terms of the truth tables that the first line applies. Since in this line Brown's and Smith's testimonies are false, Brown and Smith commit perjury in this case.
d) There is only one line (namely the third one) in which everyone's testimony is true. In this line B and S are 1 and J is 0. So, in this case Brown and Smith are innocent and Jones is guilty.
e) Line 6 in the truth table is the only line in which the innocent tells the truth and the guilty tells lies. From line 6 we read off that in this case Brown and Smith are guilty and tell lies and that Jones is innocent and tells the truth.

Solution 2.32. Let P, Q, R be the statement 'Pro wins', 'Quick wins, 'the Runners win', respectively.

P	Q	R	Trainer of Pro $R \to \neg Q$	Trainer of Quick $Q \vee R$	Trainer of Runners R
1	1	1	0	1	1
1	1	0	1	1	0
1	0	1	1	1	1
1	**0**	**0**	**1**	**0**	**0**
0	1	1	0	1	1
0	1	0	1	1	0
0	0	1	1	1	1
0	0	0	1	0	0

a) The assumption that everyone's statement is true means in terms of the truth tables that the third or seventh line applies. Assuming there is at most one winner, the third line does not apply. So, the Runners win.
b) If only the trainer of the winning club makes a true statement, Pro wins the tournament, as can be seen from the fourth line.

Solution 2.33. (a) $\neg P \wedge Q \wedge R$ (see the outline of the proof of Theorem 2.16).
(b) $(P \wedge Q \wedge R) \vee (\neg P \wedge Q \wedge R) \vee (\neg P \wedge \neg Q \wedge R)$.
(c) $P \wedge \neg P$.
(d) $\neg((P \wedge Q \wedge R) \vee (\neg P \wedge Q \wedge \neg R))$. Note: the table of $(P \wedge Q \wedge R) \vee (\neg P \wedge Q \wedge \neg R)$ corresponds with the negation of column (d).

Solution 2.34. $\neg A$ has the same truth table as $\neg A \vee \neg A$ and hence as $A \downarrow A$.
$A \vee B$ has the same truth table as $\neg(\neg A) \vee \neg(\neg B)$, hence as $\neg A \downarrow \neg B$ and therefore as $(A \downarrow A) \downarrow (B \downarrow B)$.
$A \wedge B$ has the same truth table as $\neg(\neg A \vee \neg B)$, hence as $\neg(A \downarrow B)$ and therefore as $(A \downarrow B) \downarrow (A \downarrow B)$.

Solution 2.35. i) $\wedge$ can be expressed in terms of $\vee$ and $\neg$, for $A \wedge B$ has the same truth table as $\neg(\neg A \vee \neg B)$; similarly, $\vee$ can be expressed in terms of $\wedge$ and $\neg$, for $A \vee B$ has the same truth table as $\neg(\neg A \wedge \neg B)$.
ii) $\{\to, \neg\}$ is complete, for according to Theorem 2.16 $\{\wedge, \vee, \neg\}$ is complete and both $\wedge$ and $\vee$ can be expressed in terms of $\to$ and $\neg$: $A \wedge B$ has the same truth table as $\neg(A \to \neg B)$ and $A \vee B$ has the same truth table as $(A \to B) \to B$.
$\{\to, \neg\}$ is independent, for $\to$ cannot be expressed in terms of $\neg$; more precisely, $A \to B$ does not have the same truth table as A, $\neg A$, $\neg\neg A$, B, $\neg B$ or $\neg\neg B$; and $\neg$ cannot be expressed in terms of $\to$; for suppose A is 1, then $\neg A$ is 0 and one can show that any formula, built from A and $\to$ only, is 1 if A is 1.
iii) In a similar way one shows that $\{\wedge, \neg\}$ and $\{\vee, \neg\}$ are both complete and independent.

Solution 2.36. Suppose $|$ is a binary connective such that every truthfunctional connective of (1 or) 2 arguments can be expressed in it. (*)
Then, in particular, there must be a formula A built from P and $|$ only, such that $\neg P$ has the same truth table as A (α). Now, if $1 \mid 1 = 1$, one can show that any formula, built from P and $|$ only, will have the value 1 if P is 1 (β). However, $\neg 1 = 0$ (γ). From

(α), (β) and (γ) it follows that $1 \mid 1 = 0$. In a similar way one shows that $0 \mid 0 = 1$. Consequently, the connective "|" must have one of the following four truth tables.

P	Q				
1	1	0	0	0	0
1	0	0	0	1	1
0	1	0	1	0	1
0	0	1	1	1	1

We will show next that the values of $1 \mid 0$ and $0 \mid 1$ should be the same, so that only the first and the fourth column remain and | must be either $\uparrow$ or $\downarrow$.

If $1 \mid 0 \neq 0 \mid 1$, then one can show that any formula, built from P, Q and | only, will get a different truth value if we interchange the P and the Q in it, giving P and Q the values 1 and 0 respectively (a). Under the assumption (*) there must be a formula B built from P, Q and | only, such that $P \wedge Q$ has the same truth table as B (b). However, $1 \wedge 0 = 0 \wedge 1$ (c). From (a), (b) and (c) it follows that $1 \mid 0 = 0 \mid 1$.

Solution 2.37. i) $(P \to (Q \to P)) \wedge (P \to Q \vee P)$ has the same truth table as $(\neg P \vee (\neg Q \vee P)) \wedge (\neg P \vee (Q \vee P))$.
ii) The following formulas have the same truth table:

$(P \to \neg(Q \to P)) \wedge (P \to Q \wedge P)$
$(\neg P \vee \neg(\neg Q \vee P)) \wedge (\neg P \vee (Q \wedge P))$
$(\neg P \vee (\neg\neg Q \wedge \neg P)) \wedge (\neg P \vee (Q \wedge P))$
$(\neg P \vee (Q \wedge \neg P)) \wedge (\neg P \vee (Q \wedge P))$
$(\neg P \vee Q) \wedge (\neg P \vee \neg P) \wedge (\neg P \vee Q) \wedge (\neg P \vee P)$
$(\neg P \vee Q) \wedge \neg P$

iii) $(P \to \neg(Q \to P)) \vee (P \to Q \wedge P)$ has the same truth table as:
$((\neg P \vee Q) \wedge (\neg P \vee \neg P)) \vee ((\neg P \vee Q) \wedge (\neg P \vee P))$
$((\neg P \vee Q) \wedge \neg P) \vee (\neg P \vee Q)$
$((\neg P \vee Q) \vee (\neg P \vee Q)) \wedge ((\neg P \vee Q) \vee \neg P)$
$(\neg P \vee Q)$.

Solution 2.38. a) $R \to \neg K,\ K \vdash \neg R$. The following list of formulas is a deduction of $\neg R$ from the premisses $R \to \neg K$ and K:
1. K premiss
2. $K \to (R \to K)$ axiom 1
3. $R \to K$ MP applied to 1 and 2.
4. $(R \to K) \to ((R \to \neg K) \to \neg R)$ axiom 7
5. $(R \to \neg K) \to \neg R$ MP applied to 3 and 4.
6. $R \to \neg K$ premiss
7. $\neg R$ MP applied to 5 and 6.

b) Suppose $\neg K \to R,\ K \vdash \neg R$. Then by the soundness theorem $\neg K \to R,\ K \models \neg R$. Making the truth table shows that this is false. Therefore, $\neg K \to R,\ K \nvdash \neg R$.

Solution 2.39. The following schemas are deductions of the last formula in the schema from the formulas mentioned as premisses.

(a) $$\frac{\overset{\text{premiss}}{A} \qquad \overset{\text{premiss}}{A \to B}}{B}\ MP$$

(b) $$\frac{\overset{\text{premiss}}{B} \qquad \dfrac{\overset{\text{premiss}}{A} \qquad \overset{3}{A \to (B \to A \wedge B)}}{B \to A \wedge B}\ MP}{A \wedge B}\ MP$$

(c) $$\dfrac{\overset{\text{premiss}}{A \wedge B} \qquad \overset{\text{4a}}{A \wedge B \to A}}{A}\ MP$$

(d) $$\dfrac{\overset{\text{premiss}}{A \wedge B} \qquad \overset{\text{4b}}{A \wedge B \to B}}{B}\ MP$$

(e) $$\dfrac{\overset{\text{premiss}}{A} \qquad \overset{\text{5a}}{A \to A \vee B}}{A \vee B}\ MP$$

(f) $$\dfrac{\overset{\text{premiss}}{B} \qquad \overset{\text{5b}}{B \to A \vee B}}{A \vee B}\ MP$$

(g) $$\dfrac{\overset{\text{premiss}}{\neg\neg A} \qquad \overset{\text{8}}{\neg\neg A \to A}}{A}\ MP$$

(h) $$\dfrac{\overset{\text{premiss}}{B \to A} \qquad \dfrac{\overset{\text{premiss}}{A \to B} \qquad \overset{\text{9}}{(A \to B) \to ((B \to A) \to (A \rightleftarrows B))}}{(B \to A) \to (A \rightleftarrows B)}\ MP}{A \rightleftarrows B}\ MP$$

(i) $$\dfrac{\overset{\text{premiss}}{A \rightleftarrows B} \qquad \overset{\text{10a}}{(A \rightleftarrows B) \to (A \to B)}}{A \to B}\ MP$$

(j) $$\dfrac{\overset{\text{premiss}}{A \rightleftarrows B} \qquad \overset{\text{10b}}{(A \rightleftarrows B) \to (B \to A)}}{B \to A}$$

Solution 2.40. The following list of formulas is a deduction of B from A and $\neg A$:

1.	A	premiss
2.	$A \to (\neg B \to A)$	axiom 1
3.	$\neg B \to A$	from 1 and 2 by MP
4.	$\neg A$	premiss
5.	$\neg A \to (\neg B \to \neg A)$	axiom 1
6.	$\neg B \to \neg A$	from 4 and 5 by MP
7.	$(\neg B \to A) \to ((\neg B \to \neg A) \to \neg\neg B)$	axiom 7
8.	$(\neg B \to \neg A) \to \neg\neg B$	from 3 and 7 by MP
9.	$\neg\neg B$	from 6 and 8 by MP
10.	$\neg\neg B \to B$	axiom 8
11.	B	from 9 and 10 by MP.

Solution 2.41. (a) $P \vee Q \not\models P \wedge Q$, since there is a line in the truth table in which $P \vee Q$ is 1 and $P \wedge Q$ is 0. According to the soundness theorem: if $P \vee Q \vdash P \wedge Q$, then $P \vee Q \models P \wedge Q$. Therefore, $P \vee Q \nvdash P \wedge Q$.
(b), (c) and (d) are shown in a similar way.

Solution 2.42. $P \to Q,\ P \vdash R \vee Q$. The following list of formulas is a deduction of $R \vee Q$ from P and $P \to Q$.
1. P premiss
2. $P \to Q$ premiss
3. Q MP applied to 1 and 2.
4. $Q \to R \vee Q$ axiom
5. $R \vee Q$ MP applied to 3 and 4.

Solution 2.43. $S \to H,\ \neg I \to \neg S \vdash^{?} I \to H$. Suppose this were true. Then because of the soundness theorem $S \to H,\ \neg I \to \neg S \models I \to H$. One easily checks from the truth table that this is not the case. Therefore, $S \to H,\ \neg I \to \neg S \nvdash I \to H$.

Solution 2.44. We leave the proof of (i) and (ii) to the reader. (iii) If $u(A) = 0$ and $u(A \to B) = 0$, then only the first line of the table applies; so $u(B) = 0$.
(iv) In the sixth line of the table $u(A) = 1$ and $u(B) = 2$. Hence, $u(((A \to B) \to A) \to A) = ((1 \to 2) \to 1) \to 1 = (2 \to 1) \to 1 = 0 \to 1 = 1$. If Peirce's law were generated by the production method consisting of the two logical axioms for $\to$ only, then because of (i), (ii) and (iii) $((A \to B) \to A) \to A$ would have the value 0 in every line of the table.

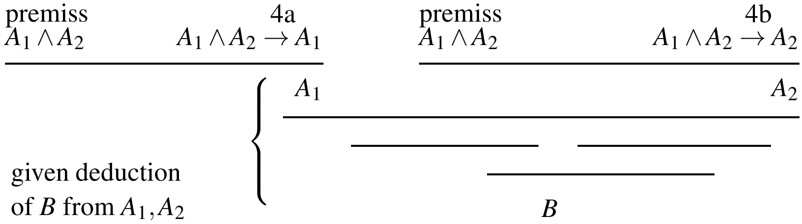
Solution 2.45. We show that $A \wedge B \to C,\ A,\ B \vdash C$. Then by two applications of the deduction theorem it follows that $A \wedge B \to C \vdash A \to (B \to C)$.

$$\dfrac{\dfrac{\overset{\text{premiss}}{B} \qquad \dfrac{\overset{\text{premiss}}{A} \quad \overset{3}{A \to (B \to A \wedge B)}}{B \to A \wedge B}}{A \wedge B} \qquad \overset{\text{premiss}}{A \wedge B \to C}}{C}$$

Solution 2.46. We show that $A \to B,\ A,\ B \to C \vdash C$. Then by three applications of the deduction theorem it follows that $\vdash (A \to B) \to (A \to ((B \to C) \to C))$.

$$\dfrac{\dfrac{\overset{\text{premiss}}{A} \qquad \overset{\text{premiss}}{A \to B}}{B} \qquad \overset{\text{premiss}}{B \to C}}{C}$$

Solution 2.47. Suppose $A_1, A_2 \vdash B$, i.e., there exists a deduction of B from A_1, A_2. We show that $A_1 \wedge A_2 \vdash B$. Then by one application of the deduction theorem it follows that $\vdash A_1 \wedge A_2 \to B$.



Solution 2.48. Suppose $\vdash (A_1 \wedge A_2) \wedge A_3 \to B$. Let (α) be a (logical) proof of $(A_1 \wedge A_2) \wedge A_3 \to B$. Then the following schema is a deduction of B from $A_1,\ A_2,\ A_3$. Note that we first deduce $(A_1 \wedge A_2) \wedge A_3$ from $A_1,\ A_2,\ A_3$ and next use $\vdash (A_1 \wedge A_2) \wedge A_3 \to B$ in order to deduce B.

$$\dfrac{(\alpha) \qquad \dfrac{\overset{\text{premiss}}{A_3} \qquad \dfrac{\dfrac{\overset{\text{premiss}}{A_2} \qquad \dfrac{\overset{\text{premiss}}{A_1} \quad \overset{3}{A_1 \to (A_2 \to A_1 \wedge A_2)}}{A_2 \to A_1 \wedge A_2}}{A_1 \wedge A_2} \qquad \overset{3}{A_1 \wedge A_2 \to (A_3 \to A_1 \wedge A_2 \wedge A_3)}}{A_3 \to A_1 \wedge A_2 \wedge A_3)}}{A_1 \wedge A_2 \wedge A_3}}{B}$$

Solution 2.49. Proof: Suppose $\vdash A \to C$ and $\vdash B \to C$. The following list of formulas is a deduction of C from $A \vee B$:
1. $A \to C$ deducible
2. $(A \to C) \to ((B \to C) \to (A \vee B \to C))$ axiom
3. $(B \to C) \to (A \vee B \to C)$ MP applied to 1 and 2.
4. $B \to C$ deducible
5. $A \vee B \to C$ MP applied to 3 and 4.
6. $A \vee B$ premiss
7. C MP applied to 5 and 6.

Solution 2.50. $A,\ B \to C \vdash A$ and $A \vdash A \vee C$; hence, $A,\ B \to C \vdash A \vee C$. (1)
$B,\ B \to C \vdash C$ and $C \vdash A \vee C$; hence, $B,\ B \to C \vdash A \vee C$. (2)
From (1) and (2), by $\vee$-elimination, $A \vee B,\ B \to C \vdash A \vee C$.

Solution 2.51. Suppose $A \vdash B$. Then, by Corollary 2.3, $A, \neg B \vdash B$.
But also $A, \neg B \vdash \neg B$. Hence, by $\neg$-introduction, $\neg B \vdash \neg A$.

Solution 2.52. $A \vdash A \vee \neg A$. Hence, by Exercise 2.51, $\neg(A \vee \neg A) \vdash \neg A$. (a)
$\neg A \vdash A \vee \neg A$. Hence, by Exercise 2.51, $\neg(A \vee \neg A) \vdash \neg(\neg A)$. (b)
From (a) and (b), by $\neg$-introduction, $\vdash \neg\neg(A \vee \neg A)$. Hence, by double negation elimination, $\vdash A \vee \neg A$.

Solution 2.53.

By weak negation elimination	$\neg A,\ \neg B,\ A \vdash C$	(1)
and	$\neg A,\ \neg B,\ B \vdash C.$	(2)
From (1) and (2), by $\vee$-elimination,	$\neg A,\ \neg B,\ A \vee B \vdash C.$	(I)
By weak negation elimination	$\neg A,\ \neg B,\ A \vdash \neg C$	(a)
and	$\neg A,\ \neg B,\ B \vdash \neg C.$	(b)
From (a) and (b), by $\vee$-elimination,	$\neg A,\ \neg B,\ A \vee B \vdash \neg C.$	(II)
From (I) and (II), by $\neg$-introduction,	$\neg A,\ \neg B \vdash \neg(A \vee B).$	

Solution 2.54. Suppose $A \vdash \neg A$. Because of $A \vdash A$, by $\neg$-introduction, $\vdash \neg A$.

Solution 2.55. Counterexample: $A = P$ (it rains) en $B = \neg P$ (it does not rain). $\models P \vee \neg P$ (it is always true that it rains or does not rain). Hence, because of the *completeness theorem*, $\vdash P \vee \neg P$.
But $\nvdash P$. For suppose $\vdash P$; then, because of the *soundness theorem*, $\models P$ (it is always true that it rains), which is false. Therefore $\nvdash P$.
Similarly, $\nvdash \neg P$. For suppose $\vdash \neg P$; then, because of the *soundness theorem*, $\models \neg P$ (it is always true that it does not rain; it never rains), which is false. Therefore, $\nvdash \neg P$.

Solution 2.56. Counterexample: $A = P$ (it rains). From the truth table we know that $\not\models P$ (it is not always true that it rains). So, because of the *soundness theorem* $\nvdash P$.
However, $\nvdash \neg P$. For suppose $\vdash \neg P$; then because of the *soundness theorem* $\models \neg P$ (it is always true that it does not rain; it never rains), which is false. Therefore, $\nvdash \neg P$.

Solution 2.57. a) Proof: Suppose $\vdash A \to B$. The following list of formulas is a deduction of B from A:

A premiss
$A \to B$ deducible
B MP applied to 1 and 2.
b) Proof: Suppose $\vdash \neg A$. Then, because of the *soundness theorem*, $\models \neg A$. (*)
We want to show that not $\vdash A$. So, suppose that $\vdash A$; then, because of the *soundness theorem*, $\models A$; but this contradicts (*). Therefore, not $\vdash A$.

Solution 2.58. We have seen in Exercise 2.40 that $A, \neg A \vdash B$. Hence, by the deduction theorem $\neg A \vdash A \to B$ (1). Also, by applying axiom 1, $B \to (A \to B)$, we know that $B \vdash A \to B$ (2). From (1) and (2) by $\vee$ elimination: $\neg A \vee B \vdash A \to B$.
a) $A \to B,\ \neg(\neg A \vee B), \neg A \vdash \neg A \vee B$ and $A \to B,\ \neg(\neg A \vee B), \neg A \vdash \neg(\neg A \vee B)$. Hence, by $\neg$-introduction, $A \to B,\ \neg(\neg A \vee B) \vdash \neg\neg A$.
b) By $\neg$-introduction, $A \to B,\ \neg(\neg A \vee B) \vdash \neg A$.

Solution 2.59. a) $A,\ B \vdash A \wedge B$, by using the axiom $A \to (B \to A \wedge B)$.
Proof of $A,\ \neg B \vdash \neg(A \wedge B)$: $A,\ \neg B,\ A \wedge B \vdash \neg B$ and $A,\ \neg B,\ A \wedge B \vdash B$ Hence, by reductio ad absurdum ($\neg$-introduction), $A,\ \neg B \vdash \neg(A \wedge B)$.
$\neg A,\ B \vdash A \vee B$ because $B \vdash A \vee B$.
$\neg A,\ \neg B \vdash \neg(A \vee B)$; see Exercise 2.53.
b) Suppose $\models E$. Then $E_1^* = E_2^* = E_3^* = E_4^* = E$.
Therefore $A,\ B \vdash E$ and $A,\ \neg B \vdash E$; hence by $\vee$-elimination: $A,\ B \vee \neg B \vdash E$.
Also $\neg A,\ B \vdash E$ and $\neg A,\ \neg B \vdash E$; hence by $\vee$-elimination: $\neg A,\ B \vee \neg B \vdash E$.
By Exercise 2.52, $\vdash B \vee \neg B$ and consequently, $A \vdash E$ and $\neg A \vdash E$.
Hence, by $\vee$-elimination: $A \vee \neg A \vdash E$ and therefore, $\vdash E$.

Solution 2.60. i)

$$\dfrac{\neg A \vee \neg B \qquad (1)\ \dfrac{{}^{3}[\neg A] \qquad \dfrac{{}^{1}[A \wedge B]}{A}\wedge E}{\neg(A \wedge B)}\neg I \qquad (2)\ \dfrac{{}^{3}[\neg B] \qquad \dfrac{{}^{2}[A \wedge B]}{B}\wedge E}{\neg(A \wedge B)}\neg I}{\neg(A \wedge B)}\vee E \quad (3)$$

ii)

$$(3)\ \dfrac{\dfrac{\neg(A \wedge B) \qquad \dfrac{\dfrac{(1)\ \dfrac{{}^{3}[\neg(\neg A \vee \neg B)] \qquad \dfrac{[\neg A]^{1}}{\neg A \vee \neg B}\vee I}{\neg\neg A}\neg I}{A}d\neg E \qquad \dfrac{(2)\ \dfrac{{}^{3}[\neg(\neg A \vee \neg B)] \qquad \dfrac{[\neg B]^{2}}{\neg A \vee \neg B}\vee I}{\neg\neg B}\neg I}{B}d\neg E}{A \wedge B}\wedge I}{\neg\neg(\neg A \vee \neg B)}\neg I}{\neg A \vee \neg B}d\neg E$$

Solution 2.61. i)

$$MP\ \frac{A \quad \overset{\text{axiom}}{A \to (B \to A)}}{B \to A} \qquad\qquad \dfrac{A \qquad (1)\ \dfrac{\dfrac{[A]^1}{B \to A}\ \to I}{A \to (B \to A)}\ \to I}{B \to A}\ \to E$$

ii)

$$\dfrac{(2)\ \dfrac{(1)\ \dfrac{\dfrac{{}^1[A] \qquad A \to B}{B} \qquad [\neg B]^2}{\neg A}\ \neg I}{}\ \to I}{\neg B \to \neg A}$$

This deduction starts as follows: $\dfrac{A \qquad A \to B}{B} \qquad \neg B$

To this corresponds: (1) $A, A \to B, \neg B \vdash B$, and
(2) $A, A \to B, \neg B \vdash \neg B$.

The deduction continues as follows: $(1)\ \dfrac{\dfrac{{}^1[A] \qquad A \to B}{B} \qquad \neg B}{\neg A}$

To this corresponds $A \to B,\ \neg B\ \vdash \neg A$, which follows from (1) and (2) by Theorem 2.25, $\neg$-introduction. And from $A \to B,\ \neg B \vdash \neg A$ it follows by Theorem 2.25, $\to$-introduction, that $A \to B \vdash \neg B \to \neg A$.

Solution 2.62. (a) $A \to B,\ A \wedge \neg B \vdash B$, and $A \to B,\ A \wedge \neg B \vdash \neg B$. Hence, by $\neg$-introduction, $A \to B \vdash \neg(A \wedge \neg B)$.
(b) The following schema is a tableau-deduction of $\neg(A \wedge \neg B)$ from $A \to B$:

$T\ A \to B,\ F\ \neg(A \wedge \neg B)$
$T\ A \to B,\ T\ A \wedge \neg B$
$T\ A \to B,\ TA,\ T\neg B$
$T\ A \to B,\ TA,\ FB$
$FA,\ TA,\ FB \mid TB,\ TA,\ FB$

(c) Suppose $A \to B$ is 1 and $\neg(A \wedge \neg B)$ is 0. Then $A \wedge \neg B$ is 1. So, A is 1 and $\neg B$ is 1. Hence, $A \to B$ is 1, A is 1 and B is 0. Then (A is 0, A is 1 and B is 0) or (B is 1, A is 1 and B is 0). Contradiction. Therefore, $A \to B \models \neg(A \wedge \neg B)$.

Solution 2.63. (a) $A \to B,\ B \to C,\ A \vdash C$. Hence, by using the deduction theorem three times, $\vdash (A \to B) \to ((B \to C) \to (A \to C))$.
(b) $F\ (A \to B) \to ((B \to C) \to (A \to C))$
$T\ A \to B,\ F\ (B \to C) \to (A \to C)$
$T\ A \to B,\ T\ B \to C,\ F\ A \to C$
$T\ A \to B,\ T\ B \to C,\ TA,\ FC$
$FA,\ T\ B \to C,\ TA,\ FC \mid TB,\ T\ B \to C,\ TA,\ FC$
$TB,\ FB,\ TA,\ FC \mid TB,\ TC,\ TA,\ FC$

(c) Suppose $(A \to B) \to ((B \to C) \to (A \to C))$ is 0. Then $A \to B$ is 1, $B \to C$ is 1, A is 1 and C is 0. So, (A is 0 and 1) or (B, $B \to C$ and A are 1 and C is 0). In the latter case, B is 1 and 0 or C is 1 and 0. Contradiction.

Solution 2.64. (a) $T\, A \to B,\ T\, \neg A \to B,\ FB$

$FA,\ T\, \neg A \to B,\ FB \mid TB,\ T\, \neg A \to B,\ FB$

$FA,\ F\neg A,\ FB \mid FA,\ TB,\ FB \mid TB,\ T\, \neg A \to B,\ FB$

$FA,\ TA,\ FB$ Note that all three tableau branches are closed.

(f) $T\, A \to B \vee C,\ F\, (A \to B) \vee (A \to C)$

$T\, A \to B \vee C,\ F\, A \to B,\ F\, A \to C$

$T\, A \to B \vee C,\ TA,\ FB,\ TA,\ FC$

$FA,\ TA,\ FB,\ TA,\ FC \mid T\, B \vee C,\ TA,\ FB,\ TA,\ FC$

$TB,\ TA,\ FB,\ TA,\ FC \mid TC,\ TA,\ FB,\ TA,\ FC$

Note that all three tableau branches are closed.

Solution 2.65. a) $R \to W, \neg R \to B, \neg B \models^{?} W$.

b) $T\, R \to W,\ T\, \neg R \to B,\ T\neg B,\ FW$

$T\, R \to W,\ T\, \neg R \to B,\ FB,\ FW$

$FR,\ T\, \neg R \to B,\ FB,\ FW \mid TW,\ T\, \neg R \to B,\ FB,\ FW$

$FR,\ F\, \neg R,\ FB,\ FW \mid FR,\ TB,\ FB,\ FW$

$FR,\ TR,\ FB,\ FW$

Note that all tableau branches are closed and hence: $R \to W, \neg R \to B, \neg B \vdash' W$.

Solution 2.66. a) $R \to \neg W, W \to H, \neg R \models^{?} H$.

b) $T\, R \to \neg W,\ T\, W \to H,\ T\neg R,\ FH$

$T\, R \to \neg W,\ T\, W \to H,\ FR,\ FH$

$T\, R \to \neg W,\ FW,\ FR,\ FH \mid T\, R \to \neg W,\ TH,\ FR,\ FH$

$FR,\ FW,\ FR,\ FH \mid T\neg W,\ FW,\ FR,\ FH$

$FW,\ FW,\ FR,\ FH$

Note that the two most left tableau branches are completed but open, i.e., not closed, while the third tableau branch is closed. From any open and completed tableau branch one read off a counterexample: give R, W and H value 0, corresponding with the occurrence of $FR,\ FW,\ FH$ in the completed open tableau branch.

R	W	H	$R \to \neg W$	$W \to H$	$\neg R$	H
0	0	0	1	1	1	0

Therefore: $R \to \neg W, W \to H, \neg R \not\models H$.

Solution 2.67. (a) The following schema is a tableau-proof of $A \to (B \to A)$:

$F\, A \to (B \to A)$

$TA,\ F\, B \to A$

$TA,\ TB,\ FA$

The other axioms are treated similarly.

(b) $A,\ A \to B \vdash' B$, for the following schema is a tableau-deduction of B from $A,\ A \to B$: $TA,\ T\, A \to B,\ FB$

$TA,\ FA,\ FB \mid TA,\ TB,\ FB.$

On the other hand, suppose $\vdash' A$ and $\vdash' A \to B$. Then there is a tableau-proof starting with FA and there is a tableau-proof starting with

$F\ A \to B$
$TA,\ FB.$
In order to show that $\vdash' B$ one has to construct a tableau-proof starting with FB.

Solution 2.68. A tableau-proof of $A \vee B$ should start with: $F\ A \vee B$
$FA,\ FB$
So, if there is a tableau-proof starting with FA or there is a tableau-proof starting with FB, then $\vdash' A \vee B$.
(b) A tableau-proof of $A \wedge B$ starts with: $F\ A \wedge B$
$FA \mid FB$
The left part is a tableau-proof of A and the right part is a tableau-proof of B.

Solution 2.69. (a) $\neg P, P \vdash Q$ (weak negation elimination). Hence, by the deduction theorem, $\neg P \vdash P \to Q$. Therefore, $\neg P,\ (P \to Q) \to P \vdash P$.
(b) $P, (P \to Q) \to P \vdash P$. So, by (a) and $\vee$-elimination, $P \vee \neg P, (P \to Q) \to P \vdash P$.
(c) By Exercise 2.52, $\vdash P \vee \neg P$. Therefore, from (b), $(P \to Q) \to P \vdash P$. So, by the deduction theorem, $\vdash ((P \to Q) \to P) \to P$.

Solution 2.70. The prisoner should reason as follows: If I wake up on Friday morning, what can I conclude. One of two things. Either they will hang me today, or else the judge was lying when he said I would hang one day this week. Suppose I somehow knew that the judge's statement that I would hang one day this week was true. Then I would know that I was to die today, and I would then know that his statement about not knowing the day of my death was false. But since I do not know that his first statement is true, I have no idea what is going to happen. Shortly before noon, they come to get him. 'Now I know', says the prisoner. 'Both statements were true'.

Let A stand for 'the prisoner will be hanged on Monday, Tuesday, Wednesday or Thursday' and B for 'the prisoner will be hanged on Friday' and let $\Box B$ stand for 'one knows B', then it is shown in Exercise 6.12 that $A \vee B,\ \Box\neg A\ \not\vdash\ \Box B$, while $\Box(A \vee B),\ \Box\neg A\ \vdash\ \Box B$ does hold. See also W.V. Quine, *On a supposed antinomy*, in *The Ways of Paradox*, and F. Norwood, *The prisoner's card game*, in *The Mathematical Intelligencer*, Vol. 4, Number 3, 1982.

Solution 2.71. This paradox is veridical if we conceive it as making clear that the promise A of the crocodile is inconsistent, more precisely $A \models B \rightleftarrows \neg B$, where B stands for 'the crocodile will eat the baby'.

Solution 2.72. Let A be the statement made by the traveller. Then the condition of the cannibals may be expressed by $(A \to B \wedge \neg R) \wedge (\neg A \to \neg B \wedge R)$, where B stands for 'the traveller will be boiled', and R for 'the traveller will be roasted'. A should be such that the truth table of the condition has only 0's and hence A should be of the form (a), (b), (c) or (d).

B	R	$(A \to B \wedge \neg R) \wedge (\neg A \to \neg B \wedge R)$	A	(a)	(b)	(c)	(d)
1	1	0	0/1	0	0	1	1
1	0	0	0	0	0	0	0
0	1	0	1	1	1	1	1
0	0	0	0/1	0	1	0	1

So, the traveller should make one of the following four statements: (a) $\neg B \wedge R$, (b) $\neg B$, (c) R, (d) $B \rightarrow R$, which has the same truth table as $\neg B \vee R$.

Solution 2.73. Similar to the barber paradox. (See Exercise 2.27.)

Solution 2.74. Similar to the barber paradox. (See Exercise 2.27.)

Solution 2.75. Let 'W' stand for 'Euathlus wins the case' and 'P' for 'Euathlus has to pay'. Then according to the contract, $W \rightarrow P$ (1) and $\neg W \rightarrow \neg P$ (2), in other words, $W \rightleftarrows P$. But according to the verdict, $W \rightarrow \neg P$ (3) and $\neg W \rightarrow P$ (4), in other words, $W \rightleftarrows \neg P$. Note that $W \rightleftarrows P$ and $W \rightleftarrows \neg P$ are inconsistent. In his argument Protagoras uses both (4) and (1), while Euathlus uses both (3) and (2) in his argument.

References

1. Austin, A.Keith, An elementary approach to *NP*-completeness; in *The American Mathematical Monthly*, vol. 90, 1983, pp. 389-399.
2. Beth, E.W., *The Foundations of Mathematics*. North-Holland Publ. Co., Amsterdam, 1959.
3. Bruijn, N.G. de, A survey of the project Automath. In: J.P. Seldin and R.J. Hindley, *Essays on Combinatory Logic, Lambda Calculus and Formalism*. Academic Press, 1980.
4. Church, A., *Introduction to Mathematical Logic*. Princeton University Press, 1956.
5. Dantzig, D. van, Siginifics and its relation to semiotics. *Library of the Tenth International Congress of Philosophy* (Amsterdam, August 11-18, 1948), Vol. 2, Philosophical Essays. Veen, Amsterdam, 1948, pp. 176-189.
6. Fitting, M., *Proof methods for modal and intuitionistic logics*. Springer. 1983.
7. Frege, G., *Begriffsschrift*. Halle, 1879.
8. Frege, G., *Begriffsschrift und andere Aufsätze*. I Angelelli (ed.), Olms, Hildesheim, 1964.
9. Gentzen, G., Untersuchungen über das logische Schliessen. *Mathematische Zeitschrift*, Vol. 39, 1934-1935, 176-210; 405-431.
10. Grice, P., *Studies in the Way of Words*. Harvard University Press, 1989.
11. Harding, S.C., Can theories be refuted? *Essays on the Duhem-Quine Thesis*. Reidel Publishing Co., Dordrecht, 1976.
12. Heijenoort, J. van, *From Frege to Gödel*. A source book in mathematical logic 1879-1931. Harvard University Press, Cambridge, Mass. 1967.
13. Jeffrey, R., *Formal Logic, its scope and limits*. McGraw-Hill, New York, 1967, 1981.
14. Kleene, S.C., *Mathematical Logic*. John Wiley and Sons Inc., New York, 1967.
15. Kleene, S.C., *Introduction to Metamathematics*. North Holland, 1962.
16. Kneale, W. and M., *The Development of Logic*. Clarendon Press, Oxford, 1962.
17. Kogel, E.A. de, Ophelders, W.M.J., A Tableaux-based Automated Theorem Prover. Appendix A in H.C.M. de Swart, *LOGIC*, Volume II, *Logic and Computer Science*, Verlag Peter Lang, Frankfurt am Main, 1994.
18. Kreisel, G. and J. Krivine, *Elements of Mathematical Logic*. North-Holland, Amsterdam, 1967.
19. Mates, B., *Stoic Logic*. University of California Press, 1953, 1973.
20. Mates, B., *Elementary Logic*. Oxford University Press, London, 1965, 1972.
21. Quine, W.V., Paradox. *Scientific American*, April 1962. Reprinted in Quine, W.V., *The Ways of Paradox and other Essays*. New York, 1966.
22. Ryle, G., Heterologicality. *Analysis*, vol. 11 (1950-51).
23. Smullyan, R.M., *First-Order Logic*. Springer Verlag, Berlin, 1968.
24. Tarski, A., Truth and Proof. *Scientific American*, June 1969, pp. 63-77.
25. Whitehead, A.N., and Russell, B., *Principia Mathematica*. Vol. 1, 1910 (2nd ed. 1925), Vol. 2, 1912 (2nd ed. 1927), Vol. 3, 1913 (2nd ed. 1927). Cambridge University Press, England.