Lorenz Halbeisen and Regula Krapf

Gödel's Theorems and Zermelo's Axioms

a firm foundation of mathematics

Sunday 14th December, 2014, 18:00

Springer

Contents

Chapter 0
A Natural Approach to Natural Numbers

In the late 19th and early 20th century, several unsuccessful attempts were made to develop the natural numbers from logic. The most promising approaches were the ones due to Frege and Russell, but also their approaches failed at the end. Even though it seems impossible to develop the natural numbers just from logic, this does not justify Kronecker's ridiculous claim that the natural numbers are given by God.

In fact, the problem with the natural numbers is, that we need the notion of finiteness in order to define them, which presuppose the existence of a kind of infinite list of objects, and it is not clear whether these objects are—in some sense—not already the natural numbers which we would like to define.

However, in our opinion there is subtle distinction between the infinite set of natural numbers and an infinite list of objects, since the set of natural numbers is an *actually infinite* set, whereas an infinite list (in contrast for example to an infinite array) is just *potentially infinite*. The difference between these two types of infinity is, that the actual infinity is something which is completed and definite and consists of infinitely many elements. On the other hand, the potential infinity—introduced by Aristotle—is something that is always finite, even though more and more elements can be added to make it arbitrarily large. For example the set of prime numbers can be considered as an actually infinite set (as Cantor did), or just as a potentially infinite list of numbers without last element which is never completed (as Euclid did).

As mentioned above, it seems that there is no way to define the natural numbers just from logic. Hence, if we would like to define them, we have to make some assumptions which cannot be formalised within logic or mathematics in general. In other words, in order to define the natural numbers we have to presuppose some *metamathematical* notions like for example the notion of F I N I T E N E S S. To emphasise this fact, we shall use a wider letter spacing for the metamathematical notions we suppose.

So, let us assume that we all have a notion of F I N I T E N E S S. Let us further assume that we have two characters, say "0" and "|". With these characters, we build now the following finite strings:

$$0 \quad | \quad || \quad ||| \quad |||| \quad ||||| \quad \ldots$$

The three dots "..." mean that we always build the next string by appending the character "|" to the string we just built. Proceeding this way, we get in fact a potentially infinite list $\mathbb{N}$ of different strings which is never completed. Thus, the list $\mathbb{N}$ is of the form

$$\mathbb{N} = [0, |, ||, |||, ||||, |||||, \ldots]$$

where each strings in the list $\mathbb{N}$ is a so-called **natural number**.

It is convenient to use arabic numbers for explicitly given natural numbers (*e.g.*, we write "1" for "|") and Latin letters like $n, m, \ldots$ for non-specified natural numbers. If n and m denote natural numbers, where n appears earlier than m in the list $\mathbb{N}$, then $n, \ldots, m$ means the natural numbers which belong to the sublist $[n, \ldots, m]$ of $\mathbb{N}$; if n appears later than m in $\mathbb{N}$, then $n, \ldots, m$ is the empty set.

We shall use natural numbers frequently as subscripts for finite lists of objects like $t_1, \ldots, t_n$. In this context we mean that for each natural number k in the list $[1, \ldots, n]$, there is an object t_k, where in the case when $n = 0$, the set of objects is empty.

If n is a natural number, then $n + 1$ denotes the natural number which appears immediately after n in the list $\mathbb{N}$; and if $n \neq 0$, then $n - 1$ denotes the natural number which appears immediately before n in the list $\mathbb{N}$.

Part I
Introduction to First-Order Logic

First-Order Logic is the system of Symbolic Logic concerned not only to represent the logical relations between sentences or propositions as wholes (like *Propositional Logic*), but also to consider their internal structure in terms of subject and predicate. First-Order Logic can be considered as a kind of language which is distinguished from higher-order languages in that it does not allow quantification over subsets of the domain of discourse or other objects of higher type (like statements of infinite length or statements about formulas). Nevertheless, First-Order Logic is strong enough to formalise all of Set Theory and thereby virtually all of Mathematics.

The goal of this brief introduction to First-Order Logic is to introduce the basic concepts of formal proofs and models, which shall be investigated further in Parts II & III.

Chapter 1
Syntax: The Grammar of Symbols

The goal of this chapter is to develop the formal language of First-Order Logic from scratch. At the same time, we introduce some terminology of the so-called meta-language, which is the language we use when we speak *about* the formal language (*e.g.*, when we like to express that two strings of symbols are equal).

Alphabet

Like any other written language, First-Order Logic is based on an *alphabet*, which consists of the following *symbols*:

(a) **Variables** such as $x, y, v_0, v_1, \ldots$, which are place holders for objects of the *domain* under consideration (which can for example be the elements of a group, natural numbers, or sets). We use mainly lower case Latin letters (with or without subscripts) for variables.
(b) **logical operators** which are "$\neg$" (*not*), "$\wedge$" (*and*), "$\vee$" (*or*), and "$\rightarrow$" (*implies*).
(c) **Logical quantifiers** which are the *existential quantifier* "$\exists$" (*there is* or *there exists*) and the *universal quantifier* "$\forall$" (*for all* or *for each*), where quantification is restricted to objects only and not to formulae or sets of objects (but the objects themselves may be sets).
(d) **Equality symbol** "$=$", which stands for the particular binary *equality relation*.
(e) **Constant symbols** like the number $\mathtt{0}$ in Peano Arithmetic, or the neutral element $\mathtt{e}$ in Group Theory. Constant symbols stand for fixed individual objects in the domain.
(f) **Function symbols** such as $\circ$ (the operation in Group Theory), or $+, \cdot, \mathtt{s}$ (the operations in Peano Arithmetic). Function symbols stand for fixed functions taking objects as arguments and returning objects as values. With each function symbol we associate a positive natural number, its co-called "arity" (*e.g.*, "$\circ$" is a 2-ary or binary function, and the successor operation "$\mathtt{s}$" is a 1-ary or unary function).

More formally, to each function symbol F we adjoin a fixed FINITE string of place holders $\mathtt{x}\cdots\mathtt{x}$ and write $F\mathtt{x}\cdots\mathtt{x}$.

(g) **Relation symbols** or **predicate constants** (such as $\in$ in Set Theory) stand for fixed relations between (or properties of) objects in the domain. Again we associate an "arity" with each relation symbol (*e.g.*, "$\in$" is a binary relation). More formally, to each relation symbol R we adjoin a fixed FINITE string of place holders $\mathtt{x}\cdots\mathtt{x}$ and write $R\mathtt{x}\cdots\mathtt{x}$.

The symbols in (a)–(d) form the core of the alphabet and are called **logical symbols**. The symbols in (e)–(g) depend on the specific topic we are investigating and are called **non-logical symbols**. The set of non-logical symbols which are used in order to formalise a certain mathematical theory is called the **language** (or **signature**) of this theory, denoted by $\mathscr{L}$, and *formulae* which are formulated in a language $\mathscr{L}$ are usually called $\mathscr{L}$-formulae. For example if we investigate groups, then the only non-logical symbols we use are "$\mathtt{e}$" and "$\circ$", thus, $\mathscr{L} = \{\mathtt{e}, \circ\}$ is the language of Group Theory.

Terms & Formulae

With the symbols of our alphabet we can now start to compose words. In the language of First-Order Logic, these words are called called *terms*.

Terms. A string of symbols is a **term**, if it results from applying FINITELY many times the following rules:

(T1) Each variable is a term.
(T2) Each constant symbol is a term.
(T3) If $\tau_1, \ldots, \tau_n$ are any terms which we have already built and $F\mathtt{x}\cdots\mathtt{x}$ is an n-ary function symbol, then $F\tau_1\cdots\tau_n$ is a term (each place holder $\mathtt{x}$ is replaced with a term).

In order to define rule (T3) we had to use variables for terms, but since the variables of our alphabet stand just for objects of the domain and not for terms or other objects of the formal language, we had to introduce new symbols. For these new symbols, which do not belong to the alphabet of the formal language, we have chosen Greek letters. In fact, we shall mainly use Greek letters for variables which stand for objects of the formal language, also to emphasise the distinction between the formal language and the metalanguage However, we shall use the Latin letters F & R as variables for function and relation symbols respectively.

To make terms, relations, and other expressions in the formal language easier to read, it is convenient to introduce some more symbols, like brackets and commas, to our alphabet. For example we usually write $F(\tau_1, \ldots, \tau_n)$ rather than $F\tau_1\cdots\tau_n$.

To some extent, terms correspond to words, since they denote objects of the domain under consideration. Like real words, they are not statements and cannot express or describe possible relations between objects. So, the next step is to build sentences, or more precisely *formulae*, with these terms.

Formulae. A string of symbols is called a **formula**, if it results from applying F I N I T E L Y many times the following rules:

(F1) If τ_1 and τ_2 are terms, then $\tau_1 = \tau_2$ is a formula.
(F2) If $\tau_1, \ldots, \tau_n$ are any terms and $R\mathsf{x}\cdots\mathsf{x}$ is any non-logical n-ary relation symbol, then $R\tau_1 \cdots \tau_n$ is a formula.
(F3) If φ is any formula which we have already built, then $\neg\varphi$ is a formula.
(F4) If φ and ψ are formulae which we have already built, then $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, and $(\varphi \rightarrow \psi)$ are formulae. (To avoid the use of brackets one could write these formulae for example in *Polish notation*, *i.e.*, $\wedge\varphi\psi$, $\vee\varphi\psi$, *et cetera*.)
(F5) If φ is a formula which we have already built, and ν is an arbitrary variable, then $\exists\nu\varphi$ and $\forall\nu\varphi$ are formulae.

Formulae of the form (F1) or (F2) are the most basic expressions we have, and since every formula is a logical connection or a quantification of these formulae, they are called **atomic formulae**.

For binary relations $R\mathsf{xx}$ it is convenient to write xRy instead of $R(x, y)$. For example we write $x \in y$ instead of $\in(x, y)$, and we write $x \notin y$ rather than $\neg(x \in y)$.

If a formula φ is of the form $\exists x\psi$ or of the form $\forall x\psi$ (for some formula ψ) and x occurs in ψ, then we say that x is in the *range* of a logical quantifier. The variable x occurring at a particular place in a formula φ is either in the range of a logical quantifier or it is not in the range of any logical quantifier. In the former case this particular instance of the variable x is **bound** in φ, and in the latter case it is **free** in φ. Notice that it is possible that a certain variable occurs in a given formula bound as well as free (*e.g.*, in $\exists z(x = z) \wedge \forall x(x = y)$, the variable x is both bound and free, whereas z is just bound and y is just free). However, one can always rename the bound variables occurring in a given formula φ such that each variable in φ is either bound or free (the rules for this procedure are given later). For a formula φ, the set of variables occurring free in φ is denoted by $\mathrm{free}(\varphi)$. A formula φ is a **sentence** (or a **closed formula**) if it contains no free variables (*i.e.*, $\mathrm{free}(\varphi) = \emptyset$). For example $\forall x(x = x)$ is a sentence but $(x = x)$ is not.

In analogy to this definition we say that a term is a **closed term** if it contains no variables. Obviously, the only terms which are closed are the constant symbols and the function symbols followed by closed terms.

Sometimes it is useful to indicate explicitly which variables occur free in a given formula φ, and for this we usually write $\varphi(x_1, \ldots, x_n)$ to indicate that $\{x_1, \ldots, x_n\} \subseteq \mathrm{free}(\varphi)$.

If φ is a formula, and τ a term, then $\varphi(x/\tau)$ is the formula we get after replacing all *free* instances of x by τ. A so-called **substitution** $\varphi(x/\tau)$ is **admissible** *iff* no free occurrence of ν in φ is in the range of a quantifier that binds any variable contained in τ (*i.e.*, for each variable ν appearing in τ, no place where ν occurs free in φ is in the range of "$\exists\nu$" or "$\forall\nu$"). For example, if $x \notin \mathrm{free}(\varphi)$, then $\varphi(x/\tau)$ is admissible for any term τ. In this case, the formulae φ and $\varphi(x/\tau)$ are identical which we express by $\varphi \equiv \varphi(x/\tau)$. In general, we use the symbol "$\equiv$" in the metalanguage to denote equality of strings of symbols of the formal language. Furthermore, if φ is a formula and the substitution $\varphi(x/\tau)$ is admissible, then we write just $\varphi(\tau)$ instead of $\varphi(x/\tau)$. To express this we write $\varphi(\tau) :\equiv \varphi(x/\tau)$, where we use "$:\equiv$" in the metalanguage to define symbols (or strings of symbols) of the formal language.

So far we have letters, and we can build words and sentences. However, these sentences are just strings of symbols without any inherent meaning. Later we shall interpret formulae in the intuitively natural way by giving the symbols the intended meaning (*e.g.*, "$\wedge$" meaning "and", "$\forall x$" meaning "for all x", *et cetera*). But before we shall do so, let us stay a little bit longer on the syntactical side—nevertheless, one should consider the formulae also from a semantical point of view.

Axioms

Below we shall label certain formulae or types of formula as **axioms**, which are used in connection with *inference rules* in order to derive further formulae. From a semantical point of view we can think of axioms as "true" statements from which we deduce or prove further results. We distinguish two types of axiom, namely *logical axioms* and *non-logical axioms* (which will be discussed later). A **logical axiom** is a sentence or formula φ which is universally valid (*i.e.*, φ is true in any possible universe, no matter how the variables, constants, *et cetera*, occurring in φ are interpreted). Usually one takes as logical axioms some minimal set of formulae that is sufficient for deriving all universally valid formulae (such a set is given below).

If a symbol is involved in an axiom which stands for an arbitrary relation, function, or even for a first-order formula, then we usually consider the statement as an **axiom schema** rather than a single axiom, since each instance of the symbol represents a single axiom. The following list of axiom schemata is a system of logical axioms.

Let φ, φ_1, φ_2, and ψ, be arbitrary first-order formulae:

L_0: $\varphi \vee \neg\varphi$,
L_1: $\varphi \to (\psi \to \varphi)$,
L_2: $(\psi \to (\varphi_1 \to \varphi_2)) \to ((\psi \to \varphi_1) \to (\psi \to \varphi_2))$,
L_3: $(\varphi \wedge \psi) \to \varphi$,
L_4: $(\varphi \wedge \psi) \to \psi$,
L_5: $\varphi \to (\psi \to (\psi \wedge \varphi))$,
L_6: $\varphi \to (\varphi \vee \psi)$,
L_7: $\psi \to (\varphi \vee \psi)$,
L_8: $(\varphi_1 \to \varphi_3) \to ((\varphi_2 \to \varphi_3) \to ((\varphi_1 \vee \varphi_2) \to \varphi_3))$,
L_9: $(\varphi \to \psi) \to ((\varphi \to \neg\psi) \to \neg\varphi)$,
L_{10}: $\neg\varphi \to (\varphi \to \psi)$.

If τ is a term, ν a variable, and the substitution $\varphi(\nu/\tau)$ is admissible, then:

L_{11}: $\forall\nu\varphi(\nu) \rightarrow \varphi(\tau)$,
L_{12}: $\varphi(\tau) \rightarrow \exists\nu\varphi(\nu)$.

If ψ is a formula and ν a variable such that $\nu \notin \mathrm{free}(\psi)$, then:

L_{13}: $\forall\nu(\psi \rightarrow \varphi(\nu)) \rightarrow (\psi \rightarrow \forall\nu\varphi(\nu))$,
L_{14}: $\forall\nu(\varphi(\nu) \rightarrow \psi) \rightarrow (\exists\nu\varphi(\nu) \rightarrow \psi)$.

What is not covered yet is the symbol "$=$", so, let us have a closer look at the binary equality relation. The defining properties of equality can already be found in Book VII, Chapter 1 of Aristotle's *Topics* [**?**], where one of the rules to decide whether two things are the same is as follows: . . . *you should look at every possible predicate of each of the two terms and at the things of which they are predicated and see whether there is any discrepancy anywhere. For anything which is predicated of the one ought also to be predicated of the other, and of anything of which the one is a predicate the other also ought to be a predicate*.

In our formal system, the binary equality relation is defined by the following three axioms.

If $\tau, \tau_1, \ldots, \tau_n, \tau_1', \ldots, \tau_n'$ are any terms, R an n-ary relation symbol (*e.g.*, the binary relation symbol "$=$"), and F an n-ary function symbol, then:

L_{15}: $\tau = \tau$,
L_{16}: $(\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n') \rightarrow (R(\tau_1, \ldots, \tau_n) \rightarrow R(\tau_1', \ldots, \tau_n'))$,
L_{17}: $(\tau_1 = \tau_1' \wedge \cdots \wedge \tau_n = \tau_n') \rightarrow (F(\tau_1, \ldots, \tau_n) = F(\tau_1', \ldots, \tau_n'))$.

Finally, we define the logical operator "$\leftrightarrow$" and the binary relation symbol "$\neq$" by stipulating

$$\varphi \leftrightarrow \psi \quad :\Longleftrightarrow \quad (\varphi \rightarrow \psi) \wedge (\psi \rightarrow \varphi)$$

$$\tau \neq \tau' \quad :\Longleftrightarrow \quad \neg(\tau = \tau')$$

where we use "$:\Longleftrightarrow$" in the metalanguage to define relations between symbols (or strings of symbols) of the formal language (*i.e.*, "$\leftrightarrow$" & "$\neq$" are just abbreviations).

This completes the list of our logical axioms. In addition to these axioms, we are allowed to state arbitrarily many formulae. In logic, such a (possibly empty) set of formulae is also called a **theory**, or, when the signature $\mathscr{L}$ is specified, an **$\mathscr{L}$-theory**. Usually, a theory consists of arbitrarily many so-called **non-logical axioms** which are sentences (*i.e.*, closed formulae). Examples of theories (*i.e.*, of sets of non-logical axioms) which will be discussed in this book are the axioms of Set Theory (see Part **??**), the axioms of *Peano Arithmetic* PA (also known as *Number Theory*), and the axioms of *Group Theory* GT, which we discuss first.

GT: The language of Group Theory is $\mathscr{L}_{\mathsf{GT}} = \{\mathsf{e}, \circ\}$, where "e" is a constant symbol and "$\circ$" is a binary function symbol.

GT_0: $\forall x\forall y\forall z(x\circ(y\circ z) = (x\circ y)\circ z)$ (*i.e.*, "$\circ$" is *associative*)
GT_1: $\forall x(\mathsf{e}\circ x = x)$ (*i.e.*, "e" is a *left-neutral* element)

GT_2: $\forall x \exists y (y \circ x = \mathtt{e})$ (*i.e.*, every element has a *left-inverse*)

PA: The language of Peano Arithmetic is $\mathscr{L}_{\mathsf{PA}} = \{\mathtt{0}, \mathtt{s}, +, \cdot\}$, where "$\mathtt{0}$" is a constant symbol, "$\mathtt{s}$" is a unary function symbol, and "$+$" & "$\cdot$" are binary function symbols.

PA_0: $\neg \exists x (\mathtt{s}x = \mathtt{0})$
PA_1: $\forall x \forall y (\mathtt{s}x = \mathtt{s}y \to x = y)$,
PA_2: $\forall x (x + \mathtt{0} = x)$
PA_3: $\forall x \forall y (x + \mathtt{s}y = \mathtt{s}(x + y))$
PA_4: $\forall x (x \cdot \mathtt{0} = \mathtt{0})$
PA_5: $\forall x \forall y (x \cdot \mathtt{s}y = (x \cdot y) + x)$

If φ is any $\mathscr{L}_{\mathsf{PA}}$-formula with $x \in \text{free}(\varphi)$, then:

PA_6: $\big(\varphi(\mathtt{0}) \wedge \forall x (\varphi(x) \to \varphi(\mathtt{s}(x)))\big) \to \forall x \varphi(x)$

Notice that PA_6 is an axiom schema, known as the **induction schema**, and not just a single axiom like PA_0–PA_5.

It is often convenient to add certain *defined symbols* to a given language so that the expressions get shorter or at least are easier to read. For example in Peano Arithmetic—which is an axiomatic system for the natural numbers—we usually replace for example the expression $\mathtt{s0}$ with $\mathtt{1}$ and $\mathtt{ss0}$ with $\mathtt{2}$. More formally, we define

$$\mathtt{1} :\equiv \mathtt{s0} \qquad \text{and} \qquad \mathtt{2} :\equiv \mathtt{ss0}\,.$$

Obviously, all that can be expressed in the language $\mathscr{L}_{\mathsf{PA}} \cup \{\mathtt{1}, \mathtt{2}\}$ can also be expressed in $\mathscr{L}_{\mathsf{PA}}$.

Formal Proofs and Tautologies

So far we have a set of logical and non-logical axioms in a certain language and can define, if we wish, as many new constants, functions, and relations as we like. However, we are still not able to deduce anything from the given axioms, since until now, we do not have *inference rules* which allow us for example to infer a certain sentence from a given set of axioms.

Surprisingly, just two **inference rules** are sufficient, namely:

$$\text{Modus Ponens } (\mathsf{MP}): \quad \frac{\varphi \to \psi,\ \varphi}{\psi} \qquad \text{and} \qquad \text{Generalisation } (\forall): \quad \frac{\varphi}{\forall \nu \varphi}.$$

In the former case we say that ψ is obtained from $\varphi \to \psi$ and φ by Modus Ponens, abbreviated (MP), and in the latter case we say that $\forall \nu \varphi$ (where ν can be any variable) is obtained from φ by Generalisation, abbreviated ($\forall$).

Using these two inference rules, we are now able to define the notion of **formal proof**: Let $\mathscr{L}$ be a signature (*i.e.*, a possibly empty set of non-logical symbols) and let T be an $\mathscr{L}$-theory (*i.e.*, a possibly empty set of $\mathscr{L}$-formulae). An $\mathscr{L}$-formula ψ is **provable** from T (or provable in T), denoted $\mathsf{T} \vdash \psi$, if there is a F I N I T E sequence $\varphi_1, \ldots, \varphi_n$ of $\mathscr{L}$-formulae such that $\varphi_n \equiv \psi$ (*i.e.*, the formulae φ_n and ψ are identical), and for all i with $1 \leqslant i \leqslant n$ we have:

- φ_i is a logical axiom, or
- $\varphi_i \in \mathsf{T}$, or
- there are $j, k < i$ such that $\varphi_j \equiv \varphi_k \to \varphi_i$, or
- there is a $j < i$ such that $\varphi_i \equiv \forall \nu\, \varphi_j$ for some variable ν.

If a formula ψ is not provable from T, *i.e.*, if there is no formal proof for ψ which uses just formulae from T, then we write $\mathsf{T} \nvdash \psi$.

Formal proofs, even of very simple statements, can get quite long and tricky. Nevertheless, we shall give two examples:

Example 1.1. For every formula φ we have:

$$\vdash \varphi \to \varphi$$

Example 1.2. $\mathsf{PA} \vdash \mathtt{s0} + \mathtt{s0} = \mathtt{ss0}$

We say that two formulae φ and ψ are **logically equivalent** (or just **equivalent**), denoted $\varphi \Leftrightarrow \psi$, if $\vdash \varphi \leftrightarrow \psi$. More formally:

$$\varphi \Leftrightarrow \psi \quad :\Longleftrightarrow \quad \vdash \varphi \leftrightarrow \psi$$

In other words, if $\varphi \Leftrightarrow \psi$, then—from a logical point of view—φ and ψ state exactly the same, and therefore we could call $\varphi \leftrightarrow \psi$ a tautology, which means *saying the same thing twice*. However, in logic, a formula φ is a **tautology** if $\vdash \varphi$. Thus, the formulae φ & ψ are equivalent if and only if $\varphi \leftrightarrow \psi$ is a tautology.

Example 1.3. For every formula φ we have:

$$\varphi \Leftrightarrow \varphi$$

In the following list we summarise some tautologies and basic facts which we shall need later.

(A.1) $\vdash \varphi \to \varphi$
(A.0) $\vdash \varphi \leftrightarrow \varphi$

(B) $\{\psi, \varphi\} \vdash \varphi \wedge \psi$

(C) $\vdash (\psi \to \varphi) \to (\psi \to \forall x\varphi)$ [for $x \notin \text{free}(\psi)$]

(D.1) $\{\varphi_0 \to \varphi_1, \varphi_1 \to \varphi_2\} \vdash \varphi_0 \to \varphi_2$
(D.2) $\{\varphi_0 \to \psi, \varphi_1 \to \psi\} \vdash (\varphi_0 \vee \varphi_1) \to \psi$
(D.3) $\{\psi \to \varphi_0, \psi \to \varphi_1\} \vdash \psi \to (\varphi_0 \wedge \varphi_1)$

(E) $\vdash \varphi \to \big(\psi \to (\varphi \wedge \psi)\big)$

(F.1) $\vdash \varphi \to \neg\neg\varphi$
(F.2) $\vdash \neg\neg\varphi \to \varphi$
(F.0) $\vdash \varphi \leftrightarrow \neg\neg\varphi$

(G.1) $\vdash (\varphi \to \psi) \to (\neg\psi \to \neg\varphi)$
(G.2) $\vdash (\neg\psi \to \neg\varphi) \to (\varphi \to \psi)$
(G.0) $\vdash (\varphi \to \psi) \leftrightarrow (\neg\psi \to \neg\varphi)$

(H.0) $\{\varphi \leftrightarrow \psi\} \vdash \neg\varphi \leftrightarrow \neg\psi$
(H.1) $\{\varphi \leftrightarrow \varphi', \psi \leftrightarrow \psi'\} \vdash (\varphi \to \psi) \leftrightarrow (\varphi' \to \psi')$
(H.2) $\{\varphi \leftrightarrow \varphi', \psi \leftrightarrow \psi'\} \vdash (\varphi \vee \psi) \leftrightarrow (\varphi' \vee \psi')$
(H.3) $\{\varphi \leftrightarrow \varphi', \psi \leftrightarrow \psi'\} \vdash (\varphi \wedge \psi) \leftrightarrow (\varphi' \wedge \psi')$

(I.1) $\vdash (\varphi_1 \wedge \varphi_2) \leftrightarrow (\varphi_2 \wedge \varphi_1)$
(I.2) $\vdash (\varphi_1 \wedge \varphi_2) \wedge \varphi_3 \leftrightarrow \varphi_1 \wedge (\varphi_2 \wedge \varphi_3)$

(J.1) $\vdash (\varphi_1 \vee \varphi_2) \leftrightarrow (\varphi_2 \vee \varphi_1)$
(J.2) $\vdash (\varphi_1 \vee \varphi_2) \vee \varphi_3 \leftrightarrow \varphi_1 \vee (\varphi_2 \vee \varphi_3)$

(K.1) $\vdash (\neg\varphi \vee \psi) \to (\varphi \to \psi)$
(K.2) $\vdash (\varphi \to \psi) \to (\neg\varphi \vee \psi)$
(K.0) $\vdash (\varphi \to \psi) \leftrightarrow (\neg\varphi \vee \psi)$

(L.1) $\vdash (\neg\varphi \vee \neg\psi) \to \neg(\varphi \wedge \psi)$
(L.2) $\vdash \neg(\varphi \wedge \psi) \to (\neg\varphi \vee \neg\psi)$
(L.0) $\vdash \neg(\varphi \wedge \psi) \leftrightarrow (\neg\varphi \vee \neg\psi)$

(M.1) $\vdash \big(\varphi_1 \to (\varphi_2 \to \varphi_3)\big) \leftrightarrow \big((\varphi_1 \wedge \varphi_2) \to \varphi_3\big)$
(M.2) $\vdash \neg(\varphi \vee \psi) \leftrightarrow (\neg\varphi \wedge \neg\psi)$

(N.1) $\vdash (\varphi_1 \wedge \varphi_2) \vee \varphi_3 \to (\varphi_1 \vee \varphi_3) \wedge (\varphi_2 \vee \varphi_3)$
(N.2) $\vdash (\varphi_1 \vee \varphi_3) \wedge (\varphi_2 \vee \varphi_3) \to (\varphi_1 \wedge \varphi_2) \vee \varphi_3$

(N.0) $\vdash (\varphi_1 \wedge \varphi_2) \vee \varphi_3 \leftrightarrow (\varphi_1 \vee \varphi_3) \wedge (\varphi_2 \vee \varphi_3)$

(O) $\vdash (\varphi_1 \vee \varphi_2) \wedge \varphi_3 \leftrightarrow (\varphi_1 \wedge \varphi_3) \vee (\varphi_2 \wedge \varphi_3)$

(P.1) $\vdash x = y \leftrightarrow y = x$
(P.2) $\vdash (x = y \wedge y = z) \rightarrow x = z$

(Q.1) $\vdash \varphi(x) \leftrightarrow \varphi(y)$ [if y does not appear in $\varphi(x)$]
(Q.2) $\vdash \exists x\varphi(x) \leftrightarrow \exists y\varphi(y)$ [if y does not appear in $\varphi(x)$]
(Q.3) $\vdash \forall x\varphi(x) \leftrightarrow \forall y\varphi(y)$ [if y does not appear in $\varphi(x)$]

(R.1) $\{\varphi \leftrightarrow \psi\} \vdash \forall x\varphi \leftrightarrow \forall x\psi$
(R.2) $\{\varphi \leftrightarrow \psi\} \vdash \exists x\varphi \leftrightarrow \exists x\psi$

(S.1) $\vdash \neg\exists x\varphi \rightarrow \forall x\neg\varphi$
(S.2) $\vdash \neg\forall x\neg\varphi \rightarrow \exists x\varphi$
(S.3) $\vdash \exists x\varphi \rightarrow \neg\forall x\neg\varphi$
(S.0) $\vdash \exists x\varphi \leftrightarrow \neg\forall x\neg\varphi$

(T) $\vdash \forall x\varphi \leftrightarrow \neg\exists x\neg\varphi$

(U.1) $\vdash \exists x\exists y\varphi \leftrightarrow \exists y\exists x\varphi$
(U.2) $\vdash \exists x\exists x\varphi \leftrightarrow \exists x\varphi$
(U.3) $\vdash \forall x\exists x\varphi \leftrightarrow \exists x\varphi$
(U.4) $\vdash \exists x\forall x\varphi \leftrightarrow \forall x\varphi$

(V.1) $\vdash \big(\exists x\varphi \wedge \exists y\psi\big) \leftrightarrow \big(\exists x\exists y(\varphi \wedge \psi)\big)$ [for $x \notin \text{free}(\psi)$, $y \notin \text{free}(\varphi)$]
(V.2) $\vdash \big(\forall x\varphi \wedge \forall y\psi\big) \leftrightarrow \big(\forall x\forall y(\varphi \wedge \psi)\big)$ [for $x \notin \text{free}(\psi)$, $y \notin \text{free}(\varphi)$]
(V.3) $\vdash \big(\exists x\varphi \wedge \forall y\psi\big) \leftrightarrow \big(\exists x\forall y(\varphi \wedge \psi)\big)$ [for $x \notin \text{free}(\psi)$, $y \notin \text{free}(\varphi)$]
(V.4) $\vdash \big(\exists x\varphi \wedge \psi\big) \leftrightarrow \big(\exists x(\varphi \wedge \psi)\big)$ [for $x \notin \text{free}(\psi)$]
(V.5) $\vdash \big(\forall x\varphi \wedge \psi\big) \leftrightarrow \big(\forall x(\varphi \wedge \psi)\big)$ [for $x \notin \text{free}(\psi)$]

(W.1) $\vdash \big(\exists x\varphi \vee \exists y\psi\big) \leftrightarrow \big(\exists x\exists y(\varphi \vee \psi)\big)$ [for $x \notin \text{free}(\psi)$, $y \notin \text{free}(\varphi)$]
(W.2) $\vdash \big(\forall x\varphi \vee \forall y\psi\big) \leftrightarrow \big(\forall x\forall y(\varphi \vee \psi)\big)$ [for $x \notin \text{free}(\psi)$, $y \notin \text{free}(\varphi)$]
(W.3) $\vdash \big(\exists x\varphi \vee \forall y\psi\big) \leftrightarrow \big(\exists x\forall y(\varphi \vee \psi)\big)$ [for $x \notin \text{free}(\psi)$, $y \notin \text{free}(\varphi)$]
(W.4) $\vdash \big(\exists x\varphi \vee \psi\big) \leftrightarrow \big(\exists x(\varphi \vee \psi)\big)$ [for $x \notin \text{free}(\psi)$]
(W.5) $\vdash \big(\forall x\varphi \vee \psi\big) \leftrightarrow \big(\forall x(\varphi \vee \psi)\big)$ [for $x \notin \text{free}(\psi)$]

Before we prove some of these tautologies and basic facts, let us first introduce a few methods and techniques which allow us to simplify formal proofs.

The Art of Proof

One of the most useful method is the so-called DEDUCTION THEOREM:

THEOREM 1.1 (DEDUCTION THEOREM). *If* T *is a theory and* $\mathsf{T} \cup \{\psi\} \vdash \varphi$*, where in the proof of* φ *from* $\mathsf{T} \cup \{\psi\}$ *the rule of* GENERALISATION *is not applied to the free variables of* ψ*, then* $\mathsf{T} \vdash \psi \to \varphi$*; and vice versa, if* $\mathsf{T} \vdash \psi \to \varphi$*, then* $\mathsf{T} \cup \{\psi\} \vdash \varphi$*:*

$$\mathsf{T} \cup \{\psi\} \vdash \varphi \quad \Longleftrightarrow \quad \mathsf{T} \vdash \psi \to \varphi \tag{DT}$$

Proof. Beweis. ⊣

Notice that the DEDUCTION THEOREM allows us under certain conditions to transform a formal proof into another. So, the DEDUCTION THEOREM is a theorem about formal proofs (*i.e.*, about sequences of formulae) and not a theorem of a theory.

As an application of the DEDUCTION THEOREM, we show that the equality relation is symmetric, which is (P.1). We first work with the empty theory and show that $\{x = y\} \vdash y = x$:

φ_1:	$(x = y \wedge x = x) \to (x = x \to y = x)$	instance of L_{17}
φ_2:	$(x = y \wedge x = x) \to x = x$	instance of L_4
φ_3:	$\varphi_1 \to (\varphi_2 \to ((x = y \wedge x = x) \to y = x))$	instance of L_2
φ_4:	$\varphi_2 \to ((x = y \wedge x = x) \to y = x)$	from φ_3 and φ_1 by MODUS PONENS
φ_5:	$(x = y \wedge x = x) \to y = x$	from φ_4 and φ_2 by MODUS PONENS
φ_6:	$x = x$	instance of L_{16}
φ_7:	$x = y$	$x = y$ belongs to $\{x = y\}$
φ_8:	$x = x \to (x = y \to (x = y \wedge x = x))$	instance of L_5
φ_9:	$x = y \to (x = y \wedge x = x)$	from φ_8 and φ_6 by MODUS PONENS
φ_{10}:	$x = y \wedge x = x$	from φ_9 and φ_7 by MODUS PONENS
φ_{11}:	$y = x$	from φ_5 and φ_{10} by MODUS PONENS

Thus, we have $\{x = y\} \vdash y = x$, and by the Deduction Theorem 1.1 we see that $\vdash x = y \to y = x$, and finally, by GENERALISATION we get

$$\vdash \forall x \forall y (x = y \to y = x).$$

We leave it as an exercise to the reader to show that the equality relation is also transitive (see EXERCISE 1).

PROPOSITION 1.2. *Let* T *be an* $\mathscr{L}$*-theory, and* $\varphi \,\&\, \psi$ *any two* $\mathscr{L}$*-formulae. Then we have:*

$$\mathsf{T} \vdash \varphi \quad \text{and} \quad \mathsf{T} \vdash \psi \quad \Longleftrightarrow \quad \mathsf{T} \vdash \varphi \wedge \psi \tag{$\wedge$}$$

Proof. First we assume $\mathsf{T} \vdash \varphi$ and $\mathsf{T} \vdash \psi$, and show $\mathsf{T} \vdash \varphi \wedge \psi$:

φ_1:	$\psi \to \big(\varphi \to (\varphi \wedge \psi)\big)$	instance of L_5
φ_2:	ψ	provable from T by assumption
φ_3:	$\varphi \to (\varphi \wedge \psi)$	from φ_1 and φ_2 by MODUS PONENS
φ_4:	φ	provable from T by assumption
φ_5:	$\varphi \wedge \psi$	from φ_3 and φ_4 by MODUS PONENS

Now we assume $\mathsf{T} \vdash \varphi \wedge \psi$, and show $\mathsf{T} \vdash \varphi$ ($\mathsf{T} \vdash \varphi$ is similar):

φ_1:	$(\varphi \wedge \psi) \to \varphi$	instance of L_3
φ_2:	$\varphi \wedge \psi$	provable from T by assumption
φ_3:	φ	from φ_1 and φ_2 by MODUS PONENS

$\dashv$

As an immediate consequence of the definition of "$\leftrightarrow$" and PROPOSITION 1.2 we get:

$$\mathsf{T} \vdash \varphi \to \psi \quad \text{and} \quad \mathsf{T} \vdash \psi \to \varphi \quad \Longleftrightarrow \quad \mathsf{T} \vdash \varphi \leftrightarrow \psi \qquad (\leftrightarrow)$$

PROPOSITION 1.3. *Let T be an $\mathscr{L}$-theory and φ an arbitrary $\mathscr{L}$-formula. Then for every $\mathscr{L}$-formula ψ we have:*

$$\mathsf{T} \vdash \varphi \wedge \neg\varphi \quad \Longrightarrow \quad \mathsf{T} \vdash \psi \qquad (\boxtimes)$$

Proof. Let ψ be any $\mathscr{L}$-formula and assume that $\mathsf{T} \vdash (\varphi \wedge \neg\varphi)$ for some $\mathscr{L}$-formula φ. We show that $\mathsf{T} \vdash \psi$:

φ_1:	$\varphi \wedge \neg\varphi$	provable from T by assumption
φ_2:	$(\varphi \wedge \neg\varphi) \to \varphi$	instance of L_3
φ_3:	φ	from φ_2 and φ_1 by MODUS PONENS
φ_4:	$(\varphi \wedge \neg\varphi) \to \neg\varphi$	instance of L_4
φ_5:	$\neg\varphi$	from φ_4 and φ_1 by MODUS PONENS
φ_6:	$\neg\varphi \to (\varphi \to \psi)$	instance of L_{10}
φ_7:	$\varphi \to \psi$	from φ_6 and φ_5 by MODUS PONENS
φ_8:	ψ	from φ_7 and φ_3 by MODUS PONENS

$\dashv$

Notice that PROPOSITION 1.3 implies that if we can derive a contradiction from T, we can derive *every* formula we like, even the impossible, which shall be denoted by

$$\mathsf{T} \vdash \boxtimes.$$

PROPOSITION 1.4 (Proof by Cases). *Let T be an $\mathscr{L}$-theory and φ, ψ, and α some $\mathscr{L}$-formulae. Then the following four statements hold:*

$$\mathsf{T} \vdash \varphi \vee \psi \ \textit{ and } \ \mathsf{T} \cup \{\varphi\} \vdash \alpha \ \textit{ and } \ \mathsf{T} \cup \{\psi\} \vdash \alpha \quad \Longrightarrow \quad \mathsf{T} \vdash \alpha \qquad (\vee 1)$$

where $(\forall)$ is not applied to any of the free variables of φ or ψ in the proof of α from $\mathsf{T} \cup \{\varphi\}$ or $\mathsf{T} \cup \{\psi\}$ respectively.

$$\mathsf{T} \cup \{\varphi\} \vdash \psi \text{ and } \mathsf{T} \cup \{\neg\varphi\} \vdash \psi \quad \Longrightarrow \quad \mathsf{T} \vdash \psi \tag{$\vee 2$}$$

where $(\forall)$ *is not applied to any of the free variables of* φ *in the proof of* ψ *from* $\mathsf{T} \cup \{\varphi\}$ *or* $\mathsf{T} \cup \{\neg\varphi\}$ *respectively.*

$$\mathsf{T} \vdash \varphi \vee \psi \quad \Longrightarrow \quad \mathsf{T} \cup \{\neg\varphi\} \vdash \psi \tag{$\vee 3$}$$

$$\mathsf{T} \vdash \varphi \vee \psi \text{ and } \mathsf{T} \cup \{\varphi\} \vdash \square \quad \Longrightarrow \quad \mathsf{T} \vdash \psi \tag{$\vee 4$}$$

Proof. Notation muss angepasst werden.

$(\vee 1)$ We assume $\mathsf{T} \vdash \varphi \vee \psi$.

$$\begin{array}{lr}
\mathsf{T} \vdash \varphi \to \alpha & \text{(DT)} \\
\quad \vdash \psi \to \alpha & \text{(DT)} \\
\quad \vdash (\varphi \to \alpha) \to ((\psi \to \alpha) \to ((\varphi \vee \psi) \to \alpha)) & (\mathrm{L}_8) \\
\quad \vdash (\psi \to \alpha) \to ((\varphi \vee \psi) \to \alpha) & \text{(MP)} \\
\quad \vdash (\varphi \vee \psi) \to \alpha & \text{(MP)} \\
\quad \vdash \varphi \vee \psi & \text{(Assumption)} \\
\quad \vdash \alpha. & \text{(MP)}
\end{array}$$

$(\vee 2)$ Is a special case of $(\vee 1)$, since $\mathsf{T} \vdash \varphi \vee \neg\varphi$ holds by (L_4).

$(\vee 3)$ We assume $\mathsf{T} \vdash \varphi \vee \psi$.

$$\begin{array}{llr}
\mathsf{T} \cup \{\neg\varphi\} & \vdash \varphi \vee \psi & \\
& \vdash \neg\varphi & \\
& \vdash (\varphi \to \psi) \to ((\psi \to \psi) \to ((\varphi \vee \psi) \to \psi)) & (\mathrm{L}_8) \\
& \vdash \neg\varphi \to (\varphi \to \psi) & (\mathrm{L}_{10}) \\
& \vdash \varphi \to \psi & \text{(MP)} \\
& \vdash (\psi \to \psi) \to ((\varphi \vee \psi) \to \psi) & \text{(MP)} \\
& \vdash (\varphi \vee \psi) \to \psi & \text{(MP)} \\
& \vdash \psi. & \text{(MP)}
\end{array}$$

$(\vee 4)$ By $(\vee 2)$ it is enough to verify $\mathsf{T} \cup \{\varphi\} \vdash \psi$ and $\mathsf{T} \cup \{\neg\varphi\} \vdash \psi$. The first statement follows directly from ($\square$) and the second one from $(\vee 3)$.

$\dashv$

COROLLARY 1.5 (Generalised Proof by Cases). *Let* T *be an* $\mathscr{L}$*-theory and* $\psi_1, \ldots, \psi_n, \varphi$ *some* $\mathscr{L}$*-formulae. Then we have:*

$$\mathsf{T} \vdash \psi_1 \vee \cdots \vee \psi_n \text{ and } \mathsf{T} \cup \{\psi_i\} \vdash \varphi \text{ for all } 1 \leqslant i \leqslant n \quad \Longrightarrow \quad \mathsf{T} \vdash \varphi,$$

where $(\forall)$ *is not applied to any of the free variables of* ψ_i *in the proof of* φ *from* $\mathsf{T} \cup \{\psi_i\}$.

Since Corollary 1.5 is just a generalization of $(\vee 1)$, we will also denote all instance of this form by $(\vee 1)$. Note that the formula $\psi_1 \vee \cdots \vee \psi_n$ is well-defined due to (J.2).

Proof of Corollary 1.5. Let $n \geqslant 2$ and assume

$$\mathsf{T} \vdash (\psi_1 \vee \cdots \vee \psi_{n-1}) \vee \psi_n \ \text{ and } \ \mathsf{T} \cup \{\psi_i\} \vdash \varphi \ \text{ for all } 1 \leqslant i \leqslant n\,.$$

Then

$$\mathsf{T} \cup \{\psi_1 \vee \cdots \vee \psi_{n-1}\} \vdash \varphi \ \text{ and } \ \mathsf{T} \cup \{\psi_n\} \vdash \varphi$$

and by $(\vee 1)$ we get $\mathsf{T} \vdash \varphi$. $\dashv$

PROPOSITION 1.6 (Contrapositon). *Let T be an $\mathscr{L}$-theory and φ & ψ two arbitrary $\mathscr{L}$-formulae. Then we have:*

$$\mathsf{T} \cup \{\neg\psi\} \vdash \neg\varphi \quad \Longrightarrow \quad \mathsf{T} \cup \{\varphi\} \vdash \psi \tag{CP}$$

Proof. By $(\vee 2)$ it suffices to show $\mathsf{T} \cup \{\varphi, \psi\} \vdash \psi$ and $\mathsf{T} \cup \{\varphi, \neg\psi\} \vdash \psi$. The first statement is obvious and the second one is a consequence of

$$\begin{aligned} \mathsf{T} \cup \{\varphi, \neg\psi\} &\vdash \neg\varphi && \\ &\vdash \varphi && \\ &\vdash \boxplus && (\wedge) \\ &\vdash \psi. && (\boxplus) \end{aligned}$$

$\dashv$

PROPOSITION 1.7 (Proof by Contradiction). *Let T be a set of formulas, and φ be an arbitrary formula. Then the statements*

$$(\lightning) \quad \mathsf{T} \cup \{\neg\varphi\} \vdash \boxplus \quad \Rightarrow \quad \mathsf{T} \vdash \varphi, \ \textit{respectively}$$

$$\mathsf{T} \cup \{\varphi\} \vdash \boxplus \quad \Rightarrow \quad \mathsf{T} \vdash \neg\varphi$$

hold, where $\boxplus :\equiv \alpha \wedge \neg\alpha$ *for any formula* α.

Proof. We consider only the first statement, since both proofs are similar. By $(\vee 2)$ it is enough to verify $\mathsf{T} \cup \{\varphi\} \vdash \varphi$ and $\mathsf{T} \cup \{\neg\varphi\} \vdash \varphi$. The first condition is clearly satisfied and the second one follows directly from $(\wedge)$ and $(\boxplus)$. $\dashv$

PROPOSITION 1.8 ($\exists-$Introduction). *Let T be a set of formulas, $\varphi(x)$ a formula with $x \in \mathrm{free}(\varphi))$ and ψ an arbitary formula. Then:*

$$(\exists) \quad \mathsf{T} \cup \{\varphi(x)\} \vdash \psi \quad \Rightarrow \quad \mathsf{T} \cup \{\exists x \varphi(x)\} \vdash \psi.$$

Proof. We will use (DT) twice:

$$\begin{aligned} T &\vdash \varphi(x) \to \psi && \text{(DT)} \\ &\vdash \forall x(\varphi(x) \to \psi) && (\forall) \\ &\vdash \forall x(\varphi(x) \to \psi) \to (\exists x \varphi(x) \to \psi) && (\mathrm{L}_{15}) \\ &\vdash \exists x \varphi(x) \to \psi. && \text{(MP)} \end{aligned}$$

⊣

THEOREM 1.9 (DEMORGAN'S LAWS). *If $\varphi_0, \dots, \varphi_n$ are formulae, then:*

(a) $\neg(\varphi_0 \wedge \cdots \wedge \varphi_n) \Leftrightarrow (\neg\varphi_1 \vee \cdots \vee \neg\varphi_n)$
(b) $\neg(\varphi_0 \vee \cdots \vee \varphi_n) \Leftrightarrow (\neg\varphi_1 \wedge \cdots \wedge \neg\varphi_n)$
(c) $\varphi_0 \to \big(\varphi_1 \to (\cdots \to \varphi_n)\cdots\big) \Leftrightarrow \neg(\varphi_0 \wedge \cdots \wedge \varphi_n)$

Proof. ⊣

THEOREM 1.10 (GENERALISED DEDUCTION THEOREM). *If T is any theory and $\mathsf{T} \cup \{\psi_1, \dots, \psi_n\} \vdash \varphi$, where in the proof of φ from $\mathsf{T} \cup \{\psi_1, \dots, \psi_n\}$ the rule of* GENERALISATION *is not applied to any of the free variables of $\psi_1, \dots, \psi_n$, then $\mathsf{T} \vdash (\psi_1 \wedge \cdots \wedge \psi_n) \to \varphi$; and vice versa:*

$$\mathsf{T} \cup \{\psi_1, \dots, \psi_n\} \vdash \varphi \quad \Longleftrightarrow \quad \mathsf{T} \vdash (\psi_1 \wedge \cdots \wedge \psi_n) \to \varphi \qquad \text{(GDT)}$$

Proof. Follows immediately from the DEDUCTION THEOREM and from part (c) of DEMORGAN'S LAWS. ⊣

THEOREM 1.11 (3-SYMBOLS). *For every each φ there is an equivalent formula ψ which contains only "$\neg$" and "$\wedge$" as logical operators and "$\exists$" as quantifier.*

Proof. ⊣

Definition of **Prenex Normal Form**, abbreviated PNF.

THEOREM 1.12 (PRENEX NORMAL FORM THEOREM). *For every formula φ there is an equivalent formula ψ which is in* PNF.

Proof. ⊣

THEOREM 1.13 (VARIABLE SUBSTITUTION THEOREM). *For every formula φ there is an equivalent formula ψ which contains just variables among $v_0, v_1, \dots$*

Proof. ⊣

Consistency & Compactness

Let T be a set of $\mathscr{L}$-formulae. We say that T is **consistent**, denoted $\mathrm{Con}(\mathsf{T})$, if there is *no* $\mathscr{L}$-formula φ such that $\mathsf{T} \vdash (\varphi \wedge \neg\varphi)$, otherwise T is called **inconsistent**, denoted $\neg\,\mathrm{Con}(\mathsf{T})$.

PROPOSITION 1.14. *Let T be a set of $\mathscr{L}$-formulae.*

(a) *If $\neg\,\mathrm{Con}(\mathsf{T})$, then for all $\mathscr{L}$-formulae ψ we have $\mathsf{T} \vdash \psi$.*
(b) *If $\mathrm{Con}(\mathsf{T})$ and $\mathsf{T} \vdash \varphi$ for some $\mathscr{L}$-formula φ, then $\mathsf{T} \nvdash \neg\varphi$.*
(c) *If $\neg\,\mathrm{Con}(\mathsf{T} + \varphi)$, for some $\mathscr{L}$-formula φ, then $\mathsf{T} \vdash \neg\varphi$.*
(d) *If $\mathsf{T} \vdash \neg\varphi$, for some $\mathscr{L}$-formula φ, then $\neg\,\mathrm{Con}(\mathsf{T} + \varphi)$.*

Proof. (a) This is just PROPOSITION 1.3.

(b) Assume that $\mathsf{T} \vdash \varphi$ and $\mathsf{T} \vdash \neg\varphi$. Then $\mathsf{T} \vdash (\varphi \wedge \neg\varphi)$, *i.e.*, $\neg\,\mathrm{Con}(\mathsf{T})$:

φ_1:	φ	provable from T by assumption
φ_2:	$\neg\varphi$	provable from T by assumption
φ_3:	$\varphi \to (\neg\varphi \to (\varphi \wedge \neg\varphi))$	instance of L_5
φ_4:	$\neg\varphi \to (\varphi \wedge \neg\varphi)$	from φ_3 and φ_1 by MODUS PONENS
φ_5:	$\varphi \wedge \neg\varphi$	from φ_4 and φ_2 by MODUS PONENS

(c) Assume that for some $\mathscr{L}$-formula φ we have $\neg\,\mathrm{Con}(\mathsf{T} + \varphi)$. By (b) we get $\mathsf{T} + \varphi \vdash \psi$ for every $\mathscr{L}$-formula ψ. In particular we get $\mathsf{T} + \varphi \vdash \neg\varphi$ and by the DEDUCTION THEOREM we get $\mathsf{T} \vdash \varphi \to \neg\varphi$:

$\mathsf{T} \vdash \varphi \to \neg\varphi$	consequence of assumption
$\mathsf{T} \vdash \varphi \to \varphi$	TAUTOLOGY (A.1)
$\mathsf{T} \vdash (\varphi \to \varphi) \to ((\varphi \to \neg\varphi) \to \neg\varphi)$	L_9
$\mathsf{T} \vdash (\varphi \to \neg\varphi) \to \neg\varphi$	by MODUS PONENS
$\mathsf{T} \vdash \neg\varphi$	by MODUS PONENS

(d) Assume that for some $\mathscr{L}$-formula φ we have $\mathsf{T} \vdash \neg\varphi$. By extending T, we also have $\mathsf{T} + \varphi \vdash \neg\varphi$:

$\mathsf{T} + \varphi \vdash \neg\varphi$	consequence of assumption
$\mathsf{T} + \varphi \vdash \varphi$	φ belongs to $\mathsf{T} + \varphi$
$\mathsf{T} + \varphi \vdash \varphi \wedge \neg\varphi$	TAUTOLOGY (B)

Hence, $\mathsf{T} + \varphi$ is inconsistent, *i.e.*, $\neg\,\mathrm{Con}(\mathsf{T} + \varphi)$. ⊣

If we design a theory T (*e.g.*, a set of axioms), we have to make sure that T is consistent. However, as we shall see later, in many cases this task is impossible.

We conclude this chapter with the COMPACTNESS THEOREM, which is a powerful tool in order to construct non-standard models of Peano Arithmetic or of Set Theory. On the one hand, it is just a consequence of the fact that formal proofs are

FINITE sequences of formulae. On the other hand, the COMPACTNESS THEOREM is the main tool to prove that a given set of sentences is consistent with some given theory.

THEOREM 1.15 (COMPACTNESS THEOREM). *Let* T *be an arbitrary set of formulae. Then* T *is consistent if and only if every finite subset* T' *of* T *is consistent.*

Proof. Obviously, if T is consistent, then every finite subset T' of T must be consistent. On the other hand, if T is inconsistent, then there is a formula φ such that $\mathsf{T} \vdash \varphi \wedge \neg\varphi$. In other words, there is a proof of $\varphi \wedge \neg\varphi$ from T. Now, since every proof is finite, there are only finitely many formulae of T involved in this proof, and if T' is this finite set of formulae, then $\mathsf{T}' \vdash \varphi \wedge \neg\varphi$, which shows that T', a finite subset of T, is inconsistent. $\dashv$

EXERCISES

0. Something with terms.

1. The equality relation is transitive.

Chapter 2
Semantics: Making Sense of the Symbols

There are two different views to a given set of formulae T, namely the *syntactical* view and the *semantical* view.

From the syntactical point of view (presented in the previous chapter), we consider the set T just as a set of well-formed formulae—regardless of their intended sense or meaning—from which we can prove some formulae. The only thing that matters is the relationship between the objects, which is given by the axioms (*i.e.*, by the formulae of T), and not the objects themselves. So, from a formal point of view there is no need to assign real objects (what ever this means) to our strings of symbols.

In contrast to this very formal syntactical view, there is also the semantical point of view from which we consider the intended meaning of the formulae in T and then seeking for a *model* in which all formulae of T become true. For this, we have to explain some basic notions of Model Theory like *structure* and *interpretation*, which we will do in an natural, informal language. In this language, we will use words like "or", "and", or phrases like "if. . .then". These words and phrases have the usual meaning. Furthermore, we assume that in our normal world, which we describe with our informal language, the basic rules of *common logic* apply. For example, a statement φ is true or false, and if φ is true, then $\neg\varphi$ is false; and vice versa. Hence, the statement "φ or $\neg\varphi$" is always true, which means that we tacitly assume the LAW OF EXCLUDED MIDDLE, also known as TERTIUM NON DATUR, which corresponds to the logical axiom L_0. Furthermore, we assume DEMORGAN'S LAWS and we apply MODUS PONENS as inference rule.

Structures & Interpretations

In order to define structures and interpretations, we have to assume some notions of NAIVE SET THEORY like *subset*, *cartesian product*, or *relation*, which shall be defined properly in Part **??**. On this occasion we also make use of the set theoretical symbol "$\in$", which stands for the binary *membership relation*.

Let $\mathscr{L}$ be an arbitrary but fixed language. An **$\mathscr{L}$-structure** $\mathbf{M}$ consists of a non-empty set A, called the **domain** of $\mathbf{M}$, together with a mapping which assigns to each constant symbol $c \in \mathscr{L}$ an element $c^{\mathbf{M}} \in A$, to each n-ary relation symbol $R \in \mathscr{L}$ a set of n-tuples $R^{\mathbf{M}}$ of elements of A, and to each n-ary function symbol $F \in \mathscr{L}$ a function $F^{\mathbf{M}}$ from n-tuples of A to A. In other word, the constant symbols become elements of A, n-ary relation symbols become subsets of A^n (*i.e.*, subsets of the n-fold cartesian product of A), and n-ary functions symbols become n-ary functions from A^n to A.

The interpretation of variables is given by a so-called assignment: An **assignment** in an $\mathscr{L}$-structure $\mathbf{M}$ is a mapping j which assigns to each variable an element of the domain A.

Finally, an **$\mathscr{L}$-interpretation** $\mathbf{I}$ is a pair $(\mathbf{M}, j)$ consisting of an $\mathscr{L}$-structure $\mathbf{M}$ and an assignment j in $\mathbf{M}$. For a variable ν, an element $a \in A$, and an assignment j in $\mathbf{M}$ we define the assignment $j\frac{a}{\nu}$ by stipulating

$$j\tfrac{a}{\nu}(\nu') = \begin{cases} a & \text{if } \nu' \equiv \nu, \\ j(\nu') & \text{otherwise.} \end{cases}$$

For an interpretation $\mathbf{I} = (\mathbf{M}, j)$ and an element $a \in A$, let

$$\mathbf{I}\tfrac{a}{\nu} := (\mathbf{M}, j\tfrac{a}{\nu}).$$

We associate with every interpretation $\mathbf{I} = (\mathbf{M}, j)$ and every $\mathscr{L}$-term τ an element $\mathbf{I}(t) \in A$ as follows:

- For a variable ν let $\mathbf{I}(\nu) := j(\nu)$.
- For a constant symbol $c \in \mathscr{L}$ let $\mathbf{I}(c) := c^{\mathbf{M}}$.
- For an n-ary function symbol $F \in \mathscr{L}$ and terms $\tau_1, \dots, \tau_n$ let

$$\mathbf{I}\big(F(\tau_1, \dots, \tau_n)\big) := F^{\mathbf{M}}\big(\mathbf{I}(\tau_1), \dots, \mathbf{I}(\tau_n)\big).$$

Now, we are able to define precisely when a formula φ becomes *true* under an interpretation $\mathbf{I} = (\mathbf{M}, j)$; in which case we write $\mathbf{I} \vDash \varphi$ and say that φ is **true** in $\mathbf{I}$ (or that φ **holds** in $\mathbf{I}$). The definition is by induction on the complexity of the formula φ. By the rules (F1)–(F5), φ must be of the form $\tau_1 = \tau_2$, $R(\tau_1, \dots, \tau_n)$, $\neg\psi$, $\psi_1 \wedge \psi_2$, $\psi_1 \vee \psi_2$, $\psi_1 \rightarrow \psi_2$, $\exists\nu\psi$, or $\forall\nu\psi$:

$$\mathbf{I} \vDash \tau_1 = \tau_2 \quad :\Longleftrightarrow \quad \mathbf{I}(\tau_1) \text{ IS THE SAME OBJECT AS } \mathbf{I}(\tau_2)$$

$$\mathbf{I} \vDash R(\tau_1, \dots, \tau_n) \quad :\Longleftrightarrow \quad \big\langle \mathbf{I}(\tau_1), \dots, \mathbf{I}(\tau_n) \big\rangle \text{ BELONGS TO } R^{\mathbf{M}}$$

$$\mathbf{I} \vDash \neg\psi \quad :\Longleftrightarrow \quad \text{NOT } \mathbf{I} \vDash \psi$$

$$\mathbf{I} \vDash \psi_1 \wedge \psi_2 \quad :\Longleftrightarrow \quad \mathbf{I} \vDash \psi_1 \text{ AND } \mathbf{I} \vDash \psi_2$$

$$\mathbf{I} \vDash \psi_1 \vee \psi_2 \quad :\Longleftrightarrow \quad \mathbf{I} \vDash \psi_1 \text{ OR } \mathbf{I} \vDash \psi_2$$

$$\mathbf{I} \vDash \psi_1 \rightarrow \psi_2 \quad :\Longleftrightarrow \quad \text{IF } \mathbf{I} \vDash \psi_1 \text{ THEN } \mathbf{I} \vDash \psi_2$$

$$\mathbf{I} \vDash \exists \nu \psi \quad :\Longleftrightarrow \quad \text{IT EXISTS } a \text{ IN } A: \ \mathbf{I}\tfrac{a}{\nu} \vDash \psi$$

$$\mathbf{I} \vDash \forall \nu \psi \quad :\Longleftrightarrow \quad \text{FOR ALL } a \text{ IN } A: \ \mathbf{I}\tfrac{a}{\nu} \vDash \psi$$

Notice that by the logical rules in our informal language, for *every* $\mathscr{L}$-formula φ we have either $\mathbf{I} \vDash \varphi$ or $\mathbf{I} \vDash \neg\varphi$. So, every $\mathscr{L}$-formula is either true or false in $\mathbf{I}$.

The following fact summarises a few immediate consequences of the definitions above:

FACT 2.1. (a) *If φ is a formula and $\nu \notin \mathrm{free}(\varphi)$, then:*

$$\mathbf{I}\tfrac{a}{\nu} \vDash \varphi \quad \textit{if and only if} \quad \mathbf{I} \vDash \varphi$$

(b) *If $\varphi(\nu)$ is a formula and the substitution $\varphi(\nu/\tau)$ is admissible, then:*

$$\mathbf{I}\tfrac{\mathbf{I}(\tau)}{\nu} \vDash \varphi(\nu) \quad \textit{if and only if} \quad \mathbf{I} \vDash \varphi(\tau)$$

Models

Let T be an arbitrary set of $\mathscr{L}$-formulae. Then an $\mathscr{L}$-structure $\mathbf{M}$ is a **model of** T if for every assignment j and for each formula $\varphi \in \mathsf{T}$ we have $(\mathbf{M}, j) \vDash \varphi$, *i.e.*, φ is true in the $\mathscr{L}$-interpretation $\mathbf{I} = (\mathbf{M}, j)$. Instead of saying "$\mathbf{M}$ is a model of T" we just write $\mathbf{M} \vDash \mathsf{T}$. If φ fails in $\mathbf{M}$, then we write $\mathbf{M} \nvDash \varphi$, which is equivalent to $\mathbf{M} \vDash \neg\varphi$, because for any $\mathscr{L}$-formula φ we have *either* $\mathbf{M} \vDash \varphi$ *or* $\mathbf{M} \vDash \neg\varphi$.

Example 2.1. Beispiel

As an immediate consequence of the definition of models we get:

FACT 2.2. *If φ is an $\mathscr{L}$-formula, ν a variable, and $\mathbf{M}$ a model of some $\mathscr{L}$-theory, then $\mathbf{M} \vDash \varphi$ if and only if $\mathbf{M} \vDash \forall\nu\varphi$.*

This leads to the following definition: Let $\langle \nu_0, \dots, \nu_n \rangle$ be the sequence of variables which appear free in the $\mathscr{L}$-formula φ, where the variables appear in the sequence as they appear in φ if one reads φ from left to right. Then the **universal closure** of φ, denoted $\overline{\varphi}$, is defined by stipulating

$$\overline{\varphi} :\equiv \forall \nu_0 \cdots \forall \nu_n \, \varphi \, .$$

As a generalisation of FACT 2.2 we get:

FACT 2.3. *If φ is an $\mathscr{L}$-formula and $\mathbf{M}$ a model of some $\mathscr{L}$-theory, then:*

$$\mathbf{M} \vDash \varphi \iff \mathbf{M} \vDash \overline{\varphi}$$

Basic Notions of Model Theory

Let $\mathscr{L}$ be a signature, *i.e.*, a possibly empty set of constant symbols c, n-ary function symbols F, and n-ary relation symbols R. Two $\mathscr{L}$-structures $\mathbf{M}$ & $\mathbf{N}$ with domains A & B are **isomorphic**, denoted $\mathbf{M} \cong \mathbf{N}$, if there is a bijection $f : A \to B$ such that

$$f\big(c^{\mathbf{M}}\big) = c^{\mathbf{N}} \quad \text{(for all } c \in \mathscr{L}\text{)}$$

and for all $a_1, \dots, a_n \in A$:

$$f\big(F^{\mathbf{M}}(a_1, \dots, a_n)\big) = F^{\mathbf{N}}\big(f(a_1), \dots, f(a_n)\big) \qquad \text{(for all } F \in \mathscr{L}\text{)}$$

$$\langle a_1, \dots, a_n \rangle \in R^{\mathbf{M}} \Leftrightarrow \big\langle f(a_1), \dots, f(a_n) \big\rangle \in R^{\mathbf{N}} \quad \text{(for all } R \in \mathscr{L}\text{)}$$

FACT 2.4. (a) *If $\mathbf{M}$ & $\mathbf{N}$ are isomorphic $\mathscr{L}$-structures and σ is an $\mathscr{L}$-sentence, then:*

$$\mathbf{M} \vDash \sigma \iff \mathbf{N} \vDash \sigma$$

(b) *If $\mathbf{M}$ & $\mathbf{N}$ are isomorphic models of some $\mathscr{L}$-theory and φ is an $\mathscr{L}$-formula, then:*

$$\mathbf{M} \vDash \varphi \iff \mathbf{N} \vDash \varphi$$

It may happen that although two $\mathscr{L}$-structures $\mathbf{M}$ & $\mathbf{N}$ are not isomorphic there is no $\mathscr{L}$-sentence that can distinguish between them. In this case we say that $\mathbf{M}$ & $\mathbf{N}$ are elementarily equivalent. More formally, we say that $\mathbf{M}$ is **elementarily equivalent** to $\mathbf{N}$, denoted $\mathbf{M} \equiv \mathbf{N}$, if each $\mathscr{L}$-sentence σ true in $\mathbf{M}$ is also true in $\mathbf{N}$. The following lemma shows that "$\equiv$" is symmetric:

LEMMA 2.5. *If $\mathbf{M}$ & $\mathbf{N}$ are $\mathscr{L}$-structures and $\mathbf{M} \equiv \mathbf{N}$, then for each $\mathscr{L}$-sentence σ:*

$$\mathbf{M} \vDash \sigma \iff \mathbf{N} \vDash \sigma$$

Proof. One direction is immediate from the definition. For the other direction, assume that σ is not true in $\mathbf{M}$, *i.e.*, $\mathbf{M} \nvDash \sigma$. Then $\mathbf{M} \vDash \neg\sigma$, which implies $\mathbf{N} \vDash \neg\sigma$, and hence, σ is not true in $\mathbf{N}$. $\dashv$

As a consequence of FACT 2.3 we get:

FACT 2.6. *If $\mathbf{M}$ & $\mathbf{N}$ are elementarily equivalent models of some $\mathscr{L}$-theory and φ is an $\mathscr{L}$-formula, then:*

$$\mathbf{M} \models \varphi \quad \Longleftrightarrow \quad \mathbf{N} \models \varphi$$

EXERCISES

2. If two structures $\mathbf{M}$ & $\mathbf{N}$ are isomorphic, then they are elementarily equivalent.
3. The converse of EXERCISE 2 does not hold.

Chapter 3
Soundness & Completeness

In this chapter we investigate the relationship between syntax and semantic. In particular, we investigate the relationship between a formal proof of a formula from a theory T and the truth-value of that formula in a model of T. In this context, two questions arise naturally:

- Is each formula φ, which is provable from some theory T, valid in every model $\mathbf{M}$ of T?
- Is every formula φ, which is valid in each model $\mathbf{M}$ of T, provable from T?

In the following section we give an answer to the former question; the answer to the latter is postponed to Part II.

Soundness Theorem

A logical calculus is called *sound*, if all what we can prove is valid (*i.e.*, true), which implies that we cannot derive a contradiction. The following theorem shows that First-Order Logic is sound.

THEOREM 3.1 (SOUNDNESS THEOREM). *Let T be a set of $\mathscr{L}$-formulae and $\mathbf{M}$ a model of T. Then for every $\mathscr{L}$-formula φ_0 we have:*

$$\mathsf{T} \vdash \varphi_0 \quad \Longrightarrow \quad \mathbf{M} \vDash \varphi_0$$

Somewhat shorter we could say:

$$\forall \varphi_0 : \ \mathsf{T} \vdash \varphi_0 \quad \Longrightarrow \quad \forall \mathbf{M}\big(\mathbf{M} \vDash \mathsf{T} \Longrightarrow \mathbf{M} \vDash \varphi_0\big)$$

Proof. First we show that all logical axioms are valid in $\mathbf{M}$. For this we have to define truth-values of composite statements in the metalanguage.
In the previous chapter we defined for example:

$$\underbrace{\mathbf{M} \vDash \varphi \wedge \psi}_{\Theta} \quad \Longleftrightarrow \quad \underbrace{\mathbf{M} \vDash \varphi}_{\Phi} \quad \text{AND} \quad \underbrace{\mathbf{M} \vDash \psi}_{\Psi}$$

Thus, in the metalanguage the statement "Θ" is true if and only if the statement "Φ AND Ψ" is true. So, the truth-value of "Θ" depends on the truth-values of "Φ" and "Ψ". In order to determine truth-values of composite statement like "Φ AND Ψ", we introduce so called *truth-tables*, in which "**1**" *stands for* "**true**" and "**0**" *stands for* "**false**":

$\boldsymbol{\Phi}$	$\boldsymbol{\Psi}$	NOT Φ	Φ AND Ψ	Φ OR Ψ	IF Φ THEN Ψ
0	0	1	0	0	1
0	1	1	0	1	1
1	0	0	0	1	0
1	1	0	1	1	1

With these truth-tables one can show that all logical axioms are valid in $\mathbf{M}$. As an example we that every instance of L_1 is valid in $\mathbf{M}$: For this, let φ_1 be an instance of L_1, *i.e.*, $\varphi_1 \equiv \varphi \to (\psi \to \varphi)$ for some $\mathscr{L}$-formulae φ & ψ. Then $\mathbf{M} \vDash \varphi_1$ *iff* $\mathbf{M} \vDash \varphi \to (\psi \to \varphi)$:

$$\underbrace{\mathbf{M} \vDash \varphi \to (\psi \to \varphi)}_{\Theta} \quad \Longleftrightarrow \quad \text{IF} \ \underbrace{\mathbf{M} \vDash \varphi}_{\Phi} \ \text{THEN} \ \underbrace{\mathbf{M} \vDash \psi \to \varphi}_{\text{IF} \ \underbrace{\mathbf{M} \vDash \psi}_{\Psi} \ \text{THEN} \ \underbrace{\mathbf{M} \vDash \varphi}_{\Phi}}$$

This shows that

$$\Theta \quad \Longleftrightarrow \quad \text{IF} \ \Phi \ \text{THEN} \ \big(\text{IF} \ \Psi \ \text{THEN} \ \Phi\big).$$

Writing the truth-table of "Θ", we see that the statement "Θ" is always true in $\mathbf{M}$:

Φ	Ψ	IF Ψ THEN Φ	IF Φ THEN (IF Ψ THEN Φ)
0	0	1	1
0	1	0	1
1	0	1	1
1	1	1	1

Therefore, $\mathbf{M} \vDash \varphi_1$, and since φ_1 was an arbitrary instance of L_1, every instance of L_1 is valid in $\mathbf{M}$.

In order to show that also the logical axioms L_{11}–L_{17} are valid in $\mathbf{M}$, we need somewhat more than just truth-tables:

Let A be the domain of $\mathbf{M}$, let j be an arbitrary assignment, and let $\mathbf{I} = (\mathbf{M}, j)$ be the corresponding $\mathscr{L}$-interpretation.

Now, we show that every instance of L_{11} is valid in $\mathbf{M}$. For this, let φ_{11} be an instance of L_{11}, *i.e.*, $\varphi_{11} \equiv \forall\nu\varphi(\nu) \to \varphi(\tau)$ for some $\mathscr{L}$-formula φ, where ν is a variable, τ a term, and the substitution $\varphi(\nu/\tau)$ is admissible. We work with $\mathbf{I}$ and show that $\mathbf{I} \vDash \varphi_{11}$.

By definition we have:

$$\mathbf{I} \vDash \forall\nu\varphi(\nu) \to \varphi(\tau) \quad \Longleftrightarrow \quad \text{IF } \mathbf{I} \vDash \forall\nu\varphi(\nu) \text{ THEN } \mathbf{I} \vDash \varphi(\tau)$$

Again by definition we have:

$$\mathbf{I} \vDash \forall\nu\varphi(\nu) \quad \Longleftrightarrow \quad \text{FOR ALL } a \text{ IN } A: \ \mathbf{I}\tfrac{a}{\nu} \vDash \varphi$$

In particular we get:

$$\mathbf{I} \vDash \forall\nu\varphi(\nu) \quad \Longrightarrow \quad \mathbf{I}\tfrac{I(\tau)}{\nu} \vDash \varphi$$

Furthermore, by FACT 2.1.(a) we get:

$$\mathbf{I} \vDash \varphi(\tau) \quad \Longleftrightarrow \quad \mathbf{I}\tfrac{I(\tau)}{\nu} \vDash \varphi(\nu)$$

Hence, we get

$$\text{IF } \mathbf{I} \vDash \forall\nu\varphi(\nu) \text{ THEN } \mathbf{I} \vDash \varphi(\tau)$$

which shows that

$$(\mathbf{M}, j) \vDash \forall\nu\varphi(\nu) \to \varphi(\tau)$$

and since the assignment j was arbitrary, we finally get:

$$\mathbf{M} \vDash \forall\nu\varphi(\nu) \to \varphi(\tau)$$

Therefore, $\mathbf{M} \vDash \varphi_{11}$, and since φ_{11} was an arbitrary instance of L_{11}, every instance of L_{11} is valid in $\mathbf{M}$.

With similar arguments one can show that also every instance of L_{12}, L_{13}, or L_{14} is valid in $\mathbf{M}$ (see EXERCISES 4–6).

Zeigen, dass auch L_{15}–L_{17} *in* $\mathbf{M}$ *gelten.*

Let now $\mathbf{M}$ be a model of T and assume that $\mathsf{T} \vdash \varphi_0$. We shall show that $\mathbf{M} \vDash \varphi_0$. For this, we notice first the following facts:

- As we have seen above, each instance of a logical axiom is valid in $\mathbf{M}$.
- Since $\mathbf{M} \vDash \mathsf{T}$, each formula of T is valid in $\mathbf{M}$.
- By the truth-tables we get

$$\text{IF } \left(\mathbf{M} \vDash \varphi \to \psi \text{ AND } \mathbf{M} \vDash \varphi \right) \text{ THEN } \mathbf{M} \vDash \psi$$

and therefore, every application of MODUS PONENS in the proof of φ_0 from T yields a valid formula (if the premisses are valid).

- Since, by FACT 2.2,

$$\mathbf{M} \vDash \varphi \iff \mathbf{M} \vDash \forall \nu \varphi(\nu)$$

every application of the GENERALISATION in the proof of φ_0 from T yields a valid formula.

From these facts it follows immediately that *each* formula in the proof of φ_0 from T is valid in $\mathbf{M}$. In particular we get

$$\mathbf{M} \vDash \varphi_0$$

which completes the proof. ⊣

The following fact summarises a few consequences of the SOUNDNESS THEOREM.

FACT 3.2.

(a) *Every tautology is valid in each model:*

$$\forall \varphi : \vdash \varphi \implies \forall \mathbf{M} : \mathbf{M} \vDash \varphi$$

(b) *If a theory* T *has a model, then* T *is consistent:*

$$\exists \mathbf{M} : \mathbf{M} \vDash \mathsf{T} \implies \mathrm{Con}(\mathsf{T})$$

(c) *The logical axioms are consistent:*

$$\mathrm{Con}(\mathsf{L}_0\text{-}\mathsf{L}_{17})$$

(d) *If a formula* φ *is not valid in* $\mathbf{M}$*, where* $\mathbf{M}$ *is a model of* T*, then* φ *is not provable from* T*:*

$$\text{IF } \left(\mathbf{M} \nvDash \varphi \text{ AND } \mathbf{M} \vDash \mathsf{T} \right) \text{ THEN } \mathsf{T} \nvdash \varphi$$

Complete Theories

An $\mathscr{L}$-theory T is called **complete**, if for every $\mathscr{L}$-sentence σ we have *either* $\mathsf{T} \vdash \sigma$ *or* $\mathsf{T} \vdash \neg\sigma$.

For an $\mathscr{L}$-theory T let $\mathbf{Th}(\mathsf{T})$ be the set of all $\mathscr{L}$-sentences σ, such that $\mathsf{T} \vdash \sigma$.

By these definitions we get that a consistent $\mathscr{L}$-theory T is complete *iff* for every $\mathscr{L}$-sentence σ we have *either* $\sigma \in \mathbf{Th}(\mathsf{T})$ *or* $\neg\sigma \in \mathbf{Th}(\mathsf{T})$.

Let $\mathbf{M}$ be a model of some $\mathscr{L}$-theory T and let $\mathsf{T}_{\mathbf{M}}$ be the set of $\mathscr{L}$-sentences σ, such that $\mathbf{M} \vDash \sigma$. Then $\mathsf{T}_{\mathbf{M}}$ is a complete theory which contains T.

Weitere Fakten und ein paar Beispiele.

EXERCISES

4. L_{12} is valid in $\mathbf{M}$.
5. L_{13} is valid in $\mathbf{M}$.
6. L_{14} is valid in $\mathbf{M}$.

Part II
Gödel's Completeness Theorem

In this part of the book we shall prove Gödel's COMPLETENESS THEOREM and show several consequences.

Gödel proved his famous theorem in his doctoral dissertation *Über die Vollständigkeit des Logikkalküls* [**?**] which was completed in 1929. In 1930, he published the same material as in the doctoral dissertation in a rewritten and shortened form in [**?**]. However, instead of presenting Gödel's original proof we decided to follow Henkin's construction, which can be found in [**?**] (see also [**?**]), since it fits better in the logical framework developed in Part I. Even though Henkin's construction works also for uncountable signatures, we shall prove in Chapter 6 COMPLETENESS THEOREM with an ultraproduct construction, using ŁOŠ'S THEOREM.

We would like to mention that in our proof of the COMPLETENESS THEOREM for countable signatures (carried out in Chapters 4 & 5), one only has to assume the existence of *potentially infinite* sets but no instance of an *actually infinite* set is required (see also Chapter 0).

Chapter 4
Maximally Consistent Extensions

Throughout this chapter we require that all formulae are written in Polish notation and that the variables are among $v_0, v_1, v_2, \ldots$ Recall that by the PRENEX NORMAL FORM THEOREM 1.12 and by the VARIABLE SUBSTITUTION THEOREM 1.13, every formula can be transformed into an equivalent formula of the required form.

Maximally Consistent Theories

Let $\mathscr{L}$ be an arbitrary signature and let T be an $\mathscr{L}$-theory. We say that T is **maximally consistent** if T is consistent and for every $\mathscr{L}$-sentence σ we have *either* $\sigma \in \mathsf{T}$ *or* $\neg \operatorname{Con}(\mathsf{T} + \sigma)$. In other words, a consistent theory T is maximally consistent if no proper extension of T is consistent.

The following fact is just a reformulation of the definition.

FACT 4.1. *Let $\mathscr{L}$ be a signature and let T be a consistent $\mathscr{L}$-theory. Then T is maximally consistent iff for every $\mathscr{L}$-sentence σ, either $\sigma \in \mathsf{T}$ or $\mathsf{T} \vdash \neg\sigma$.*

Proof. By THEOREM 1.14.(c)&(d) we have:

$$\neg \operatorname{Con}(\mathsf{T} + \sigma) \quad \Longleftrightarrow \quad \mathsf{T} \vdash \neg\sigma$$

Hence, an $\mathscr{L}$-theory is maximally consistent *iff* for every $\mathscr{L}$-sentence σ, either $\sigma \in \mathsf{T}$ or $\mathsf{T} \vdash \neg\sigma$. ⊣

As a consequence of FACT 4.1 we get

LEMMA 4.2. *Let $\mathscr{L}$ be a signature and let T be a consistent $\mathscr{L}$-theory. Then T is maximally consistent iff for every $\mathscr{L}$-sentence σ, either $\sigma \in \mathsf{T}$ or $\neg\sigma \in \mathsf{T}$.*

Proof. We have to show that the following equivalence holds:

$$\forall \sigma \big(\sigma \in \mathsf{T} \;\; or \;\; \mathsf{T} \vdash \neg\sigma\big) \quad \Longleftrightarrow \quad \forall \sigma \big(\sigma \in \mathsf{T} \;\; or \;\; \neg\sigma \in \mathsf{T}\big)$$

($\Rightarrow$) Assume that for every $\mathscr{L}$-sentence σ we have $\sigma \in \mathsf{T}$ or $\mathsf{T} \vdash \neg\sigma$. If $\sigma \in \mathsf{T}$, then the implication obviously holds. If $\sigma \notin \mathsf{T}$, then $\mathsf{T} \vdash \neg\sigma$, and since T is consistent, this implies $\mathsf{T} \nvdash \sigma$. Now, by TAUTOLOGY (F.0), this implies $\mathsf{T} \nvdash \neg\neg\sigma$ and by our assumption we finally get $\neg\sigma \in \mathsf{T}$.

($\Leftarrow$) Assume that for every $\mathscr{L}$-sentence σ we have $\sigma \in \mathsf{T}$ or $\neg\sigma \in \mathsf{T}$. If $\sigma \in \mathsf{T}$, then the implication obviously holds. Now, if $\sigma \notin \mathsf{T}$, then by our assumption we have $\neg\sigma \in \mathsf{T}$, which obviously implies $\mathsf{T} \vdash \neg\sigma$. ⊣

Maximally consistent theories have similar features as complete theories: Recall that an $\mathscr{L}$-theory T is complete if for every $\mathscr{L}$-sentence σ we have *either* $\mathsf{T} \vdash \sigma$ *or* $\mathsf{T} \vdash \neg\sigma$.

As an immediate consequence of the definitions we get

FACT 4.3. *Let $\mathscr{L}$ be a signature, let T be a consistent $\mathscr{L}$-theory, and let $\mathbf{Th}(\mathsf{T})$ be the set of all $\mathscr{L}$-sentences which are provable from T.*

(a) *If T is complete, then $\mathbf{Th}(\mathsf{T})$ is maximally consistent.*
(b) *If T is maximally consistent, then $\mathbf{Th}(\mathsf{T})$ is equal to T.*

The next lemma gives a condition under which a theory can be extended to maximally consistent theory.

LEMMA 4.4. *If an $\mathscr{L}$-theory T has a model, then T has a maximally consistent extension.*

Proof. Let $\mathbf{M}$ be a model of the $\mathscr{L}$-theory T and let $\mathsf{T}_{\mathbf{M}}$ be the set of $\mathscr{L}$-sentences σ such that $\mathbf{M} \vDash \sigma$. Then $\mathsf{T}_{\mathbf{M}}$ is obviously a maximally consistent theory which contains T. ⊣

Later we shall see that every consistent theory has a model. For this, we first show how a consistent theory can be extended to a maximally consistent theory.

Universal List of Sentences

Let $\mathscr{L}$ be an arbitrary but fixed countable signature, where by "countable" we mean that the symbols in $\mathscr{L}$ can be listed in a FINITE or POTENTIALLY INFINITE list $L_{\mathscr{L}}$.

First, we encode the symbols of $\mathscr{L}$ corresponding to the order in which they appear in the list $L_{\mathscr{L}}$: The first symbol is encoded with "2", the second with "22", the third with "222", and so on. For every symbol $\zeta \in L_{\mathscr{L}}$ let $\#\zeta$ denote the code of ζ. So, the code of a symbol of $\mathscr{L}$ is just a sequence of 2's.

Furthermore, we encode the logical symbols as follows:

Symbol ζ	Code $\#\zeta$
$=$	11
$\neg$	1111
$\wedge$	111111
$\vee$	11111111
$\rightarrow$	1111111111
$\exists$	111111111111
$\forall$	11111111111111
v_0	1
v_1	111
$\vdots$	$\vdots$
v_n	$\underbrace{1111\ \ldots\ 11111}_{(2n+1)\ 1\text{'s}}$

In the next step, we encode strings of symbols: Let $\bar{\zeta} \equiv \zeta_1\zeta_2\zeta_3\ldots\zeta_n$ be a finite string of symbols, then

$$\#\bar{\zeta} := \#\zeta_1 0 \#\zeta_2 0 \#\zeta_3 \ldots 0 \#\zeta_n$$

For a string $\#\zeta$ (*i.e.*, a string of 0's, 1's, and 2's) let $|\#\zeta|$ be the length of $\#\zeta$ (*i.e.*, the number of 0's, 1's, and 2's which appear in $\#\zeta$).

Now, we order the codes of strings of symbols by their length and lexicographically, where $0 < 1 < 2$. If, with respect to this ordering, $\#\zeta_1$ is less than $\#\zeta_2$, we write $\zeta_1 \prec \zeta_2$.

Finally, let $\Lambda_{\mathscr{L}} = [\sigma_1, \sigma_2, \ldots]$ be the potentially infinite list of all $\mathscr{L}$-sentences, ordered by "$\prec$" (*i.e.*, $\sigma_i \prec \sigma_j$ *iff* $i < j$). We call $\Lambda_{\mathscr{L}}$ the **universal list of $\mathscr{L}$-sentences**.

Lindenbaum's Lemma

In this section we show that every consistent set of $\mathscr{L}$-sentences T can be extended to a maximally consistent set of $\mathscr{L}$-sentences $\overline{\mathsf{T}}$. Since the universal list of $\mathscr{L}$-sentences contains all possible $\mathscr{L}$-sentences, every set of $\mathscr{L}$-sentences can be can be listed in a (finite or potentially infinite) list. So, we do not have to assume that the (possibly infinite) set of $\mathscr{L}$-sentences T is completed and definite.

LINDENBAUM'S LEMMA 4.5. *Let $\mathscr{L}$ be a countable signature and let T be a consistent set of $\mathscr{L}$-sentences. Furthermore, let σ_0 be an $\mathscr{L}$-sentences which cannot* *be proved from* T, *i.e.*, $\mathsf{T} \nvdash \sigma_0$. *Then there exists a maximally consistent set* $\overline{\mathsf{T}}$ *of* $\mathscr{L}$*-sentences which contains* $\neg\sigma_0$ *as well as all the sentences of* T.

Proof. Let $\Lambda_{\mathscr{L}} = [\sigma_1, \sigma_2, \ldots]$ be the universal list of $\mathscr{L}$-sentences. First we extend $\Lambda_{\mathscr{L}}$ with the $\mathscr{L}$-sentence $\neg\sigma_0$; let $\Lambda^0_{\mathscr{L}} = [\neg\sigma_0, \sigma_1, \sigma_2, \ldots]$.

Now, we go through the list $\Lambda^0_{\mathscr{L}}$ and define step by step a list $\overline{\mathsf{T}}$ of $\mathscr{L}$-sentences: For this, we define T_0 as the empty list, *i.e.*, $T_0 := [\,]$. If T_n is already defined, then

$$T_{n+1} := \begin{cases} T_n + [\sigma_n] & \text{if } \operatorname{Con}(T + T_n + \sigma_n), \\ T_n & \text{otherwise.} \end{cases}$$

Let $\overline{\mathsf{T}} = [\sigma_{i_0}, \sigma_{i_1}, \ldots]$ be the resulting list, *i.e.*, $\overline{\mathsf{T}}$ is the union of all the T_n's.

Notice that the construction only works if we assume the LAW OF EXCLUDED MIDDLE: Even in the case when we cannot decide whether $T + T_n + \sigma_n$ is consistent or not, we assume, from a metamathematical point of view, that *either* $T + T_n + \sigma_n$ is consistent *or* $T + T_n + \sigma_n$ is inconsistent (and neither both nor none).

CLAIM. $\overline{\mathsf{T}}$ *is a maximally consistent set of* $\mathscr{L}$*-sentences which contains* $\neg\sigma_0$ *as well as all the sentences of* T.

Proof of Claim. First we show that $\neg\sigma_0$ belongs to $\overline{\mathsf{T}}$, then we show that $\mathsf{T} + \overline{\mathsf{T}}$ is consistent (which implies that $\overline{\mathsf{T}}$ is consistent), in a third step we show that $\overline{\mathsf{T}}$ contains T, and finally we show that for every $\mathscr{L}$-sentence σ we have either $\sigma \in \overline{\mathsf{T}}$ or $\neg \operatorname{Con}(\overline{\mathsf{T}} + \sigma)$.

$\neg\sigma_0$ *belongs to* $\overline{\mathsf{T}}$: Since $\mathsf{T} \nvdash \sigma_0$, by PROPOSITION 1.14.(c) we have $\operatorname{Con}(\mathsf{T} + \neg\sigma_0)$, and since $T_0 = [\,]$, we also have $\operatorname{Con}(\mathsf{T} + T_0 + \neg\sigma_0)$. Thus, $\neg\sigma_0 \in T_1$ (in fact $T_1 = [\neg\sigma_0]$) which shows that $\neg\sigma_0 \in \overline{\mathsf{T}}$.

$\mathsf{T} + \overline{\mathsf{T}}$ *is consistent*: By the COMPACTNESS THEOREM 1.15 it is enough to show that every finite subset of $\mathsf{T} + \overline{\mathsf{T}}$ is consistent. So, let $\mathsf{T}' + T_k$ be a finite subset of $\mathsf{T} + \overline{\mathsf{T}}$, where T' is a finite subset of T and T_k is some finite initial segment of the list $\overline{\mathsf{T}}$. Notice that since $\mathsf{T} + \neg\sigma_0$ is consistent, also $\mathsf{T}' + \neg\sigma_0$ is consistent. If $T_k = [\,]$ or $T_k = [\neg\sigma_0]$, this implies that also $\mathsf{T}' + T_k$ is consistent. Otherwise, $T_k = [\ldots, \sigma_n]$ for some σ_n in $\Lambda^0_{\mathscr{L}}$, which implies that $T_k = T_n + [\sigma_n]$. Now, by construction we get $\operatorname{Con}(\mathsf{T} + T_n + \sigma_n)$, which implies the consistency of $\mathsf{T}' + T_k$.

$\overline{\mathsf{T}}$ *contains all sentences of* T: For every $\sigma \in \mathsf{T}$ there is a $\sigma_n \in \Lambda^0_{\mathscr{L}}$ such that $\sigma \equiv \sigma_n$. By $\operatorname{Con}(\mathsf{T} + T_n + \sigma_n)$ we get $\sigma_n \in T_{n+1}$, hence, $\sigma_n \in \overline{\mathsf{T}}$ and therefore $\sigma \in \overline{\mathsf{T}}$.

For every σ, *either* $\sigma \in \overline{\mathsf{T}}$ *or* $\neg \operatorname{Con}(\overline{\mathsf{T}} + \sigma)$: For every $\mathscr{L}$-sentence σ there is a $\sigma_n \in \Lambda^0_{\mathscr{L}}$ such that $\sigma \equiv \sigma_n$. By the law of excluded middle, we have *either* $\operatorname{Con}(\mathsf{T} + T_n + \sigma_n)$, which implies $\sigma_n \in T_{n+1}$ and therefore $\sigma \in \overline{\mathsf{T}}$, *or* $\neg \operatorname{Con}(\mathsf{T} + T_n + \sigma_n)$, which implies $\neg \operatorname{Con}(\overline{\mathsf{T}} + \sigma_n)$, *i.e.*, $\neg \operatorname{Con}(\overline{\mathsf{T}} + \sigma)$. $\dashv_{\text{Claim}}$

Thus, the list $\overline{\mathsf{T}}$ has all the required properties, which completes the proof. $\dashv$

The following fact summarises the main properties of $\overline{\mathsf{T}}$.

FACT 4.6. *Let* $\mathsf{T}, \overline{\mathsf{T}}$, *and* σ_0 *be as above, and let* σ *and* σ' *be any* $\mathscr{L}$*-sentences.*

(a) $\neg\sigma_0 \in \overline{\mathsf{T}}$.

(b) *Either* $\sigma \in \overline{\mathsf{T}}$ *or* $\neg\sigma \in \overline{\mathsf{T}}$.

(c) *If* $\mathsf{T} \vdash \sigma$, *then* $\sigma \in \overline{\mathsf{T}}$.

(d) $\overline{\mathsf{T}} \vdash \sigma$ *iff* $\sigma \in \overline{\mathsf{T}}$.

(e) *If* $\sigma \Leftrightarrow \sigma'$, *then* $\sigma \in \overline{\mathsf{T}}$ *iff* $\sigma' \in \overline{\mathsf{T}}$.

Proof. (a) follows by construction of $\overline{\mathsf{T}}$.

Since $\overline{\mathsf{T}}$ is maximally consistent, (b) follows by LEMMA 4.2.

For (c), notice that $\mathsf{T} \vdash \sigma$ implies $\neg\,\mathrm{Con}(\mathsf{T} + \neg\sigma)$, hence $\neg\sigma \notin \overline{\mathsf{T}}$ and by (b) we get $\sigma \in \overline{\mathsf{T}}$.

For (d), let us first assume $\overline{\mathsf{T}} \vdash \sigma$. This implies $\mathrm{Con}(\overline{\mathsf{T}} + \sigma)$, hence $\mathrm{Con}(\mathsf{T} + \sigma)$, and by construction of $\overline{\mathsf{T}}$ we get $\sigma \in \overline{\mathsf{T}}$. On the other hand, if $\sigma \in \overline{\mathsf{T}}$, then we obviously have $\overline{\mathsf{T}} \vdash \sigma$.

For (e), recall that $\sigma \Leftrightarrow \sigma'$ is just an abbreviation for $\vdash \sigma \leftrightarrow \sigma'$. Thus, (e) follows immediately from (d). ⊣

Of course, this can work out only when the $\mathscr{L}$-sentences in $\overline{\mathsf{T}}$ "behave" like valid sentences in a model, which is indeed the case—as the following proposition shows.

PROPOSITION 4.7. *Let* $\overline{\mathsf{T}}$ *be as above, and let* $\sigma, \sigma_1, \sigma_2$ *be any* $\mathscr{L}$*-sentences.*

(a) $\neg\sigma \in \overline{\mathsf{T}} \iff$ NOT $\sigma \in \overline{\mathsf{T}}$

(b) $\wedge\sigma_1\sigma_2 \in \overline{\mathsf{T}} \iff \sigma_1 \in \overline{\mathsf{T}}$ AND $\sigma_2 \in \overline{\mathsf{T}}$

(c) $\vee\sigma_1\sigma_2 \in \overline{\mathsf{T}} \iff \sigma_1 \in \overline{\mathsf{T}}$ OR $\sigma_2 \in \overline{\mathsf{T}}$

(d) $\rightarrow\sigma_1\sigma_2 \in \overline{\mathsf{T}} \iff$ IF $\sigma_1 \in \overline{\mathsf{T}}$ THEN $\sigma_2 \in \overline{\mathsf{T}}$

Proof. (a) Follows immediately from FACT 4.6.(b).

(b) First notice that by FACT 4.6.(d), $\wedge\sigma_1\sigma_2 \in \overline{\mathsf{T}}$ *iff* $\overline{\mathsf{T}} \vdash \wedge\sigma_1\sigma_2$. Thus, by L_3 & L_4 and (MP) we get $\overline{\mathsf{T}} \vdash \sigma_1$ and $\overline{\mathsf{T}} \vdash \sigma_2$. Thus, by FACT 4.6.(d), we get $\sigma_1 \in \overline{\mathsf{T}}$ AND $\sigma_2 \in \overline{\mathsf{T}}$. On the other hand, if $\sigma_1 \in \overline{\mathsf{T}}$ AND $\sigma_2 \in \overline{\mathsf{T}}$, then, by FACT 4.6.(d), we get $\overline{\mathsf{T}} \vdash \sigma_1$ and $\overline{\mathsf{T}} \vdash \sigma_2$. Now, by TAUTOLOGY (B), this implies $\overline{\mathsf{T}} \vdash \wedge\sigma_1\sigma_2$, and by by FACT 4.6.(d) we finally get $\wedge\sigma_1\sigma_2 \in \overline{\mathsf{T}}$.

(c) & (d) follow from FACT 4.6.(e) and the fact that for each formula σ there is an equivalent formula σ' which contains neither "$\vee$" nor "$\rightarrow$" (see THEOREM **??**). ⊣

EXERCISES

7. Show that all the logical axioms of propositional logic (*i.e.*, L_0–L_{10}) were used in the proofs of FACT 4.1, LEMMA 4.2, FACT 4.6, and PROPOSITION 4.7. Notice that in the proof of FACT 4.1, we used THEOREM 1.14.(c)&(d)

Chapter 5
Models of Countable Theories

As in the previous chapter, we require that all formulae are written in Polish notation and that the variables are among $v_0, v_1, v_2, \ldots$ Furthermore, let $\mathscr{L}$ be a countable signature, let T be a consistent $\mathscr{L}$-theory, and let σ_0 be an $\mathscr{L}$-sentence which is not provable from T. Finally, let $\overline{\mathsf{T}}$ be the maximally consistent extension of $\mathsf{T} + \neg\sigma_0$ as above.

We shall now construct a model of $\overline{\mathsf{T}}$. For this, we first extend the signature $\mathscr{L}$ by adding some new constant symbols, then we extend the theory $\overline{\mathsf{T}}$, and finally we construct the model.

Extending the Language

A string of symbols is a **term-constant**, if it results from applying F I N I T E L Y many times the following rules:

(C0) Each closed (*i.e.*, variable-free) $\mathscr{L}$-term is a term-constant.
(C1) If $\tau_0, \ldots, \tau_{n-1}$ are any term-constants which we have already built and F is an n-ary function symbol, then $F\tau_0 \cdots \tau_{n-1}$ is a term-constant.
(C2) For any natural numbers i, n, if $\tau_0, \ldots, \tau_{n-1}$ are any term-constants which we have already built, then $(i, \tau_0, \ldots, \tau_{n-1}, n)$ is a term-constant.

The strings $(i, \tau_0, \ldots, \tau_{n-1}, n)$ which are built with rule (C2) are called **special constants**. Notice that for $n = 0$, $(i, \tau_0, \ldots, \tau_{n-1}, n)$ becomes $(i, 0)$.

Let $\mathscr{L}_c$ be the signature $\mathscr{L}$ extended with the countably many special constants. In order to write the special constants in a list, we first encode them and then define an ordering on the codes.

First we encode closed $\mathscr{L}$-terms as above with strings of $\mathsf{0}$'s and $\mathsf{2}$'s. Now, let $c^{\bar{\tau}}_{i,n} \equiv (i, \tau_0, \ldots, \tau_{n-1}, n)$ be a special constant, where the codes of $\tau_0, \ldots, \tau_{n-1}$ are already defined. Then we encode $c^{\bar{\tau}}_{i,n}$ as follows:

$$\begin{array}{rcccccccccccc} c^{\bar\tau}_{i,n} & \equiv & (& \boldsymbol{i} & \boldsymbol{,} & \boldsymbol{\tau_0} & \boldsymbol{,} & \dots & \boldsymbol{,} & \boldsymbol{\tau_{n-1}} & \boldsymbol{,} & \boldsymbol{n} &) \\ & & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow \\ \#c^{\bar\tau}_{i,n} & \equiv & \mathsf{6} & \underbrace{\mathsf{1}\dots\mathsf{1}}_{i\text{-times } \mathsf{1}} & \mathsf{8} & \#\tau_0 & \mathsf{8} & \dots & \mathsf{8} & \#\tau_{n-1} & \mathsf{8} & \underbrace{\mathsf{1}\dots\mathsf{1}}_{n\text{-times } \mathsf{1}} & \mathsf{9} \end{array}$$

The codes of special constants are ordered by their length and lexicographically, where $\mathsf{0} < \mathsf{1} < \ldots < \mathsf{8} < \mathsf{9}$.

Finally, let $\Lambda_c = [c_0, c_1, \ldots]$ be the potentially infinite list of all special constants, ordered with respect to the ordering of their codes.

Extending the Theory

In this section we shall add witnesses for certain existential $\mathscr{L}_c$-sentences σ_i in the list $\overline{\mathsf{T}} = [\sigma_0, \sigma_1, \ldots, \sigma_i, \ldots]$, where an $\mathscr{L}_c$-sentence is existential if it is of the form $\exists\nu\varphi$. The witnesses we choose from the list Λ_c of special constants. In order to make sure that we have a witness for each existential $\mathscr{L}_c$-sentence (and not just for $\mathscr{L}$-sentences), and also to make sure that the choice of witnesses do not lead to a contradiction, we have to choose the witnesses carefully. For this we introducte the following notion: An $\mathscr{L}$-sentence $\sigma_i \in \overline{\mathsf{T}}$ is in **special prenex normal form**, denoted sPNF, if σ_i is in PNF and

$$\sigma_i \equiv \forall\!\!\!/_0 v_0 \, \forall\!\!\!/_1 v_1 \ldots \forall\!\!\!/_n v_n \sigma_{i,n}$$

where each $\forall\!\!\!/_m$ (for $0 \leqslant m \leqslant n$) stands for either "$\exists$" or "$\forall$", $\sigma_{i,n}$ is quantifier free, and each variable $v_0, \ldots, v_n$ appears free in $\sigma_{i,n}$. Notice that by the PRENEX NORMAL FORM THEOREM 1.12 and the VARIABLE SUBSTITUTION THEOREM 1.13, for every $\mathscr{L}$-sentence σ there is an equivalent $\mathscr{L}$-sentence σ' which is in sPNF.

Let $\sigma_i \in \overline{\mathsf{T}}$ and let $c^{\bar t}_{i,n} \equiv (i, t_0, \ldots, t_{n-1}, n)$ be a special constant. Then we say that $\boldsymbol{c^{\bar t}_{i,n}}$ **witnesses** $\boldsymbol{\sigma_i}$ if:

- σ_i is in sPNF,
- "$\exists v_n$" appears in σ_i, and
- for all $m < n$: if "$\exists v_m$" appears in σ_i, then $t_m \equiv (i, t_0, \ldots, t_{m-1}, m)$.

If an $\mathscr{L}$-sentence $\sigma_i \in \overline{\mathsf{T}}$ is in sPNF and "$\exists v_n$" or "$\forall v_n$" appear in σ_i, then

$$\sigma_i \equiv \forall\!\!\!/_0 v_0 \, \forall\!\!\!/_1 v_1 \cdots \forall\!\!\!/_n v_n \sigma_{i,n}(v_0, \ldots, v_n)$$

where $\sigma_{i,n}(v_0, \ldots, v_n)$ is an $\mathscr{L}$-formula in which each variable $v_0, \ldots, v_n$ appears free.

Now, we go through the list $\Lambda_c = [c_0, c_1, \ldots]$ of special constants and extend step by step the list $\overline{\mathsf{T}} = [\sigma_0, \sigma_1, \ldots]$: For this, we first stipulate $T_0 := \overline{\mathsf{T}}$. If T_j is already defined and that $c_j \equiv (i, t_0, \ldots, t_{n-1}, n)$. We have the following two cases:

Case 1. The special constant c_j does not witness the $\mathscr{L}$-sentence $\sigma_i \in \overline{\mathsf{T}}$. In this case we set $T_{j+1} := T_j$.

Case 2. The special constant c_j witnesses $\sigma_i \in \overline{\mathsf{T}}$. In this case we insert the $\mathscr{L}_c$-sentence

$$\sigma_{i,n}[c_j] \equiv \sigma_{i,n}(v_0/t_0, \ldots, v_{n-1}/t_{n-1}, v_n/c_j)$$

into the list T_j on the place which corresponds to the code $\#\sigma_{i,n}[c_j]$. The extended list is then T_{j+1}.

Finally, let $\overline{\mathsf{T}}_c$ be the resulting list, *i.e.*, $\overline{\mathsf{T}}_c$ is the union of all the T_j's.

LEMMA 5.0. *$\overline{\mathsf{T}}_c$ is consistent.*

Proof. By construction of $\overline{\mathsf{T}}$ we have $\mathrm{Con}(\overline{\mathsf{T}})$. Now, assume towards a contradiction that $\mathrm{Con}(\overline{\mathsf{T}}_c)$ is inconsistent. Then, by the COMPACTNESS THEOREM 1.15, we find finitely many $\mathscr{L}_c$-sentences $\sigma_{i,n}[c_j]$ in $\overline{\mathsf{T}}_c$ such that

$$\neg\,\mathrm{Con}\left(\overline{\mathsf{T}} + \left\{\sigma_{i_1,n_1}[c_{j_1}], \ldots, \sigma_{i_k,n_k}[c_{j_k}]\right\}\right).$$

Without loss of generality we may assume that $\sigma_{i_1,n_1}[c_{j_1}], \ldots, \sigma_{i_k,n_k}[c_{j_k}]$ are such that the sum $n_1 + \ldots + n_k + k$ is minimal.

Now, for term-constants τ we define the height $h(\tau)$ as follows: If τ is a closed $\mathscr{L}$-term, then $h(\tau) := 0$. If $\tau_0, \ldots, \tau_{n-1}$ are term-constants und $F \in \mathscr{L}$ is an n-ary function symbol, then

$$h(F\tau_0 \cdots \tau_{n-1}) := \max\left\{h(\tau_0), \ldots, h(\tau_{n-1})\right\}.$$

Finally, if $\tau \equiv (i, \tau_0, \ldots, \tau_{n-1}, n)$ is a special constant, then

$$h(\tau) := 1 + \max\left\{h(\tau_0), \ldots, h(\tau_{n-1})\right\}.$$

Without loss of generality we may assume that $h(c_{j_k}) = \max\left\{h(c_{j_1}, \ldots, h(c_{j_k})\right\}$. To simplify the notation, let $\Sigma := \left\{\sigma_{i_1,n_1}[c_{j_1}], \ldots, \sigma_{i_k,n_{k-1}}[c_{j_{k-1}}]\right\}$; furthermore we write i, n, j instead of i_k, n_k, j_k respectively.

Now, we consider the $\mathscr{L}_c$-sentence $\sigma_{i,n}[c_j]$. For this, let $c_j \equiv (i, t_0, \ldots, t_{n-1}, n)$, *i.e.*,

$$\sigma_{i,n}[c_j] \equiv \sigma_{i,n}(v_0/t_0, \ldots, v_{n-1}/t_{n-1}, v_n/c_j).$$

Since c_j witnesses σ_i, "$\exists v_n$" appears in σ_i, *i.e.*,

$$\sigma_{i,n-1}(v_0, \ldots, v_{n-1}) \equiv \exists v_n \sigma_{i,n}(v_0, \ldots, v_{n-1}, v_n).$$

To simplify the notation again we set

$$\tilde{\sigma}(v_n) :\equiv \sigma_{i,n}(v_0/t_0, \ldots, v_{n-1}/t_{n-1}, v_n).$$

CLAIM. $\neg\,\mathrm{Con}\left(\overline{\mathsf{T}} + \Sigma + \sigma_{i,n}[c_j]\right) \implies \neg\,\mathrm{Con}\left(\overline{\mathsf{T}} + \Sigma + \exists v_n \tilde{\sigma}(v_n)\right)$

Proof of Claim. If $\overline{\mathsf{T}} + \Sigma + \sigma_{i,n}[c_j]$ is inconsistent, then $\overline{\mathsf{T}} + \Sigma + \sigma_{i,n}[c_j] \vdash \boxtimes$ and with the DEDUCTION THEOREM we get

$$\overline{\mathsf{T}} + \Sigma \vdash \sigma_{i,n}[c_j] \rightarrow \boxtimes .$$

In the latter proof we replace the special constant c_j throughout the proof with a variable ν which does not occur, neither in the proof nor in $\sigma_{i,n}$. Notice that every logical axiom becomes a logical axiom of the same type and that $\mathscr{L}$-sentences of $\overline{\mathsf{T}}$ are not affected (which do not contain any of the special constants). Furthermore, also $\mathscr{L}_c$-sentences of Σ are not affected since they do not contain the special constant c_j (otherwise, the height $h(c_j)$ would not be maximal). Finally, each application of MODUS PONENS or GENERALISATION becomes a new application of the same inference rule (notice that we do not apply GENERALISATION to ν, since otherwise, we would have applied GENERALISATION to c_j, but c_j is a constant). It follows that we obtain a proof of $\tilde{\sigma}(\nu) \rightarrow \boxtimes$ from $\overline{\mathsf{T}} + \Sigma$:

$\overline{\mathsf{T}} + \Sigma \vdash \tilde{\sigma}(\nu) \rightarrow \boxtimes$	by construction
$\overline{\mathsf{T}} + \Sigma \vdash \forall\nu\big(\tilde{\sigma}(\nu) \rightarrow \boxtimes\big)$	by GENERALISATION
$\overline{\mathsf{T}} + \Sigma \vdash \forall\nu\big(\tilde{\sigma}(\nu) \rightarrow \boxtimes\big) \rightarrow \big(\exists\nu\tilde{\sigma}(\nu) \rightarrow \boxtimes\big)$	L_{14}
$\overline{\mathsf{T}} + \Sigma \vdash \exists\nu\tilde{\sigma}(\nu) \rightarrow \boxtimes$	by MODUS PONENS
$\overline{\mathsf{T}} + \Sigma \vdash \exists v_n\tilde{\sigma}(v_n) \rightarrow \boxtimes$	TAUTOLOGY (Q.2)

Therefore, we finally have $\neg\,\mathrm{Con}\,\big(\overline{\mathsf{T}} + \Sigma + \exists v_n\tilde{\sigma}(v_n)\big)$. $\dashv_{\text{Claim}}$

We now write again i_k, n_k, j_k instead of i, n, j respectively and consider the following three cases:

Case 1. If $n_k = 0$, then $\sigma_{i_k} \equiv \exists v_0\tilde{\sigma}$, *i.e.*, $\neg\,\mathrm{Con}\,\big(\overline{\mathsf{T}} + \Sigma\big)$. So,

$$\neg\,\mathrm{Con}\,\Big(\overline{\mathsf{T}} + \big\{\sigma_{i_1,n_1}[c_{j_1}], \ldots, \sigma_{i_{k-1},n_{k-1}}[c_{j_{k-1}}]\big\}\Big)$$

which is a contradiction to the minimality of $n_1 + \ldots + n_k + k$ (*i.e.*, the choice of $\sigma_{i_1,n_1}[c_{j_1}], \ldots, \sigma_{i_k,n_k}[c_{j_k}]$), since

$$n_1 + \ldots + n_{k-1} + (k-1) < n_1 + \ldots + n_k + k\,.$$

Case 2. If $n_k > 0$ and "$\exists v_m$" appears in σ_{i_k} for some $m < n_k$, then

$$\mathrm{Con}\,\Big(\overline{\mathsf{T}} + \Sigma + \sigma_{i_k,m}(v_0/t_0, \ldots v_m/t_m)\Big)\,.$$

Otherwise, we would have

$$n_1 + \ldots + n_{k-1} + m + k < n_1 + \ldots + n_k + k$$

which is again a contradiction to the minimality of $n_1 + \ldots + n_k + k$.

Case 3. If, for some $m + 1 < n_k$, we have

$$\operatorname{Con}\left(\overline{\mathsf{T}} + \Sigma + \sigma_{i_k,m}(v_0/t_0, \dots v_m/t_m)\right)$$

and "$\forall v_{m+1}$" appears in σ_{i_k}, then, by L_{11}, we get

$$\operatorname{Con}\left(\overline{\mathsf{T}} + \Sigma + \sigma_{i_k,m}(v_0/t_0, \dots v_m/t_m, v_{m+1}/t_{m+1})\right).$$

Combining the Cases 1–3 we get that $\overline{\mathsf{T}} + \Sigma + \sigma_{i_k}[c_{\underline{j_k}}]$ is consistent, which contradicts our primary assumption. Hence, the $\mathscr{L}_c$-theory $\overline{\mathsf{T}}_c$ is consistent. ⊣

Completeness Theorem for Countable Signatures

In this section we shall construct a model of the $\mathscr{L}_c$-theory $\overline{\mathsf{T}}_c$, which is of course also a model of the $\mathscr{L}$-theory $\mathsf{T} + \neg\sigma_0$. However, since we extended the signature $\mathscr{L}$, we first have to extend the binary relation "$=$" as well as relation symbols in $\mathscr{L}$ to the new closed $\mathscr{L}_c$-terms.

LEMMA 5.1. *The list $\overline{\mathsf{T}}_c$ can be extended to a consistent list $\widetilde{\mathsf{T}}$ of $\mathscr{L}_c$-sentence, such that the new $\mathscr{L}_c$-sentences are variable-free and for each variable-free $\mathscr{L}_c$-senctence σ we have*

$$\textit{either} \quad \sigma \in \widetilde{\mathsf{T}} \quad \textit{or} \quad \neg\sigma \in \widetilde{\mathsf{T}}.$$

Proof. Like in the proof of LINDENBAUM'S LEMMA 4.5, we go through the list of all variable-free $\mathscr{L}_c$-sentences and successively extend the list $\overline{\mathsf{T}}_c$ to a maximally consistent list $\widetilde{\mathsf{T}}$. ⊣

Now we are ready to construct the domain of a model of $\widetilde{\mathsf{T}}$, which shall be a list of lists: For this, let

$$\Lambda_\tau = [t_0, t_1, \dots, t_n, \dots]$$

be the list of all term-constants (ordered with respect to the encoding above). We go through the list Λ_τ and construct step by step a list of lists: First, we set $A_0 := \big[[\,]\big]$. Now, assume that A_n is already defined. Then consider the $\mathscr{L}_c$-sentences

$$t_n = t_0, t_n = t_1, \dots, t_n = t_{n-1}.$$

If $t_n = t_m$ is one of these sentences and $t_n = t_m$ belongs to $\widetilde{\mathsf{T}}$, then we append t_n to the list in A_n which contains t_m; the resulting list is A_{n+1}. If none of the sentences $t_n = t_m$ belongs to $\widetilde{\mathsf{T}}$, then $A_{n+1} := A_n + \big[[t_n]\big]$.

Let $A = \big[[t_{n_0}, \dots], [t_{n_1}, \dots] \dots\big]$ be the resulting list, *i.e.*, A is the union of all the A_n's.

The lists in the list A is the domain of our model $\mathbf{M}$ of $\widetilde{\mathsf{T}}$. In order to simplify the notation, for term-constants τ let $\tilde{\tau}$ be the unique list of A which contains τ.

In order to get an $\mathscr{L}_c$-structure $\mathbf{M}$ with domain A, we have to define a mapping which assigns to each constant symbol $c \in \mathscr{L}_c$ an element $c^{\mathbf{M}} \in A$, to each n-ary function symbol $F \in \mathscr{L}$ a function $F^{\mathbf{M}} : A^n \to A$, and to each n-ary relation symbol $R \in \mathscr{L}$ a set $R^{\mathbf{M}} \subseteq A^n$:

- If $c \in \mathscr{L}_c$ is a constant symbol of $\mathscr{L}$ or a special constant, then let
$$c^{\mathbf{M}} := \widetilde{c}\,.$$
- If $F \in \mathscr{L}$ is an n-ary function symbol and $\widetilde{t}_1, \dots, \widetilde{t}_n$ are elements of A, then let
$$F^{\mathbf{M}}\widetilde{t}_1 \cdots \widetilde{t}_n := \widetilde{Ft_1 \cdots t_n}\,.$$
- If $R \in \mathscr{L}$ is an n-ary relation symbol and $\widetilde{t}_1, \dots, \widetilde{t}_n$ are elements of A, then we define
$$\langle \widetilde{t}_1, \dots, \widetilde{t}_n \rangle \in R^{\mathbf{M}} \quad :\Longleftrightarrow \quad Rt_1 \cdots t_n \in \widetilde{\mathsf{T}}\,.$$

FACT 5.2. *The definitions above, which rely on representatives of the lists in A, are well-defined.*

Proof. This follows easily by L_{15}–L_{17} and the construction of $\widetilde{\mathsf{T}}$; the details are left as an exercise to the reader. ⊣

THEOREM 5.3. *The $\mathscr{L}_c$-structure $\mathbf{M}$ is a model of $\widetilde{\mathsf{T}}$, and consequently also of $\mathsf{T} + \neg\sigma_0$.*

Proof. We have to show that for each $\mathscr{L}_c$-sentence $\sigma \in \widetilde{\mathsf{T}}$, $\mathbf{M} \vDash \sigma$, *i.e.*,
$$\forall \sigma\big(\sigma \in \widetilde{\mathsf{T}} \implies \mathbf{M} \vDash \sigma\big)\,.$$

First notice that for $\mathscr{L}_c$-sentences $\sigma \,\&\, \sigma'$ with $\sigma \Leftrightarrow \sigma'$ (*i.e.*, $\vdash \sigma \leftrightarrow \sigma'$), by the SOUNDNESS THEOREM **??** we get
$$\mathbf{M} \vDash \sigma \quad \Longleftrightarrow \quad \mathbf{M} \vDash \sigma'\,.$$

So, by the 3-SYMBOLS THEOREM 1.11 it is enough to prove the theorem only for $\mathscr{L}_c$-sentences σ which are either atomic or of the form $\neg\sigma'$, $\wedge\sigma_1\sigma_2$, or $\exists\nu\sigma'$.

We first consider the case when σ is variable-free. By LEMMA 5.1 we know that for each variable-free $\mathscr{L}_c$-sentences σ we have either $\sigma \in \widetilde{\mathsf{T}}$ or $\neg\sigma \in \widetilde{\mathsf{T}}$. Thus, we must show that for these sentences we have
$$\sigma \in \widetilde{\mathsf{T}} \quad \Longleftrightarrow \quad \mathbf{M} \vDash \sigma\,.$$

If σ is atomic, then either $\sigma \equiv t_1 = t_2$ (for some term-constants $t_1 \,\&\, t_1$) or $\sigma \equiv Rt_1 \cdots t_n$ (for term-constants $t_1, \dots, t_n$ and an n-ary relation symbol $R \in \mathscr{L}$), and by construction of $\mathbf{M}$ we get $\sigma \in \widetilde{\mathsf{T}} \Longleftrightarrow \mathbf{M} \vDash \sigma$.

Now, assume towards a contradiction that there exists a variable-free $\mathscr{L}_c$-sentence σ_0 such that either $\sigma_0 \in \widetilde{\mathsf{T}}$ *and* $\mathbf{M} \nvDash \sigma_0$, or $\sigma_0 \notin \widetilde{\mathsf{T}}$ *and* $\mathbf{M} \vDash \sigma_0$. Without loss of generality we may assume that σ_0 has as few logical symbols as possible. Notice that we already know that σ_0 is not atomic. We consider the following cases:

$\sigma_0 \equiv \neg\sigma$: Since σ has less logical symbols than σ_0, we have $\sigma \in \widetilde{\mathsf{T}}$ if and only if $\mathbf{M} \vDash \sigma$. This shows that

$$\neg\sigma \notin \widetilde{\mathsf{T}} \iff \mathbf{M} \nvDash \neg\sigma$$

which is a contradiction to the choice of σ_0.

$\sigma_0 \equiv \wedge\sigma_1\sigma_2$: Since σ_1 as well as σ_2 has less logical symbols than σ_0, we have $\sigma_1 \in \widetilde{\mathsf{T}}$ if and only if $\mathbf{M} \vDash \sigma_1$, as well as $\sigma_2 \in \widetilde{\mathsf{T}}$ if and only if $\mathbf{M} \vDash \sigma_2$. This shows that

$$\wedge\sigma_1\sigma_2 \in \widetilde{\mathsf{T}} \iff \mathbf{M} \vDash \wedge\sigma_1\sigma_2$$

which is a contradiction to the choice of σ_0.

Now, we consider the case when σ contains variables and show that for every $\sigma \in \widetilde{\mathsf{T}}$ we have $\mathbf{M} \vDash \sigma$; If σ is an $\mathscr{L}_c$-sentence which belongs to $\widetilde{\mathsf{T}}$, then there exists a $\sigma' \in \overline{\mathsf{T}}_c$ in sPNF such that $\sigma \Leftrightarrow \sigma'$. In particular we get $\mathbf{M} \vDash \sigma$ if and only if $\mathbf{M} \vDash \sigma'$.

Assume towards a contradiction that there is an $\mathscr{L}_c$-sentence $\sigma' \in \overline{\mathsf{T}}_c$ in sPNF for which we have $\mathbf{M} \nvDash \sigma'$. Notice that since $\sigma' \in \overline{\mathsf{T}}_c$, we have $\sigma' \in \widetilde{\mathsf{T}}$, in particular we get $\widetilde{\mathsf{T}} \vdash \sigma'$. For σ' there are natural numbers i, m, n with $m < n$ and term-constants $t_0, \ldots, t_{m-1}$, such that

$$\sigma' \equiv \forall\!\!\!\exists_m v_m \cdots \forall\!\!\!\exists_n v_n \sigma_{i,m}(v_0/t_0, \ldots, v_{m-1}/t_{m-1}, v_m, \ldots, v_n)\,,$$

where each $\forall\!\!\!\exists_k$ (for $m \leqslant k \leqslant n$) stands for either "$\exists$" or "$\forall$" and $\sigma_{i,n}$ is quantifier free.

Because $\mathbf{M} \nvDash \sigma'$, we get $\mathbf{M} \vDash \neg\sigma'$, and for $\neg\sigma'$ we have:

$$\neg\sigma' \equiv \overline{\forall\!\!\!\exists}_m v_m \cdots \overline{\forall\!\!\!\exists}_n v_n \neg\sigma_{i,n}(v_0/t_0, \ldots, v_{m-1}/t_{m-1}, v_m, \ldots, v_n)$$

where for $m \leqslant k \leqslant n$, the quantifier $\overline{\forall\!\!\!\exists}_k$ is "$\exists$" if $\forall\!\!\!\exists_k$ is "$\forall$", and vice versa.

For each k with $m \leqslant k \leqslant n$, we replace in $\sigma_{i,n}$ step by step the variable v_k with a term-constant t_k as follows:

- If $\forall\!\!\!\exists_k$ is the quantifier "$\forall$", then

$$\mathbf{M} \vDash \exists v_k \cdots \neg\sigma_{i,n}(v_0/t_0, \ldots, v_k, \ldots)\,.$$

Hence, there exists a $\tilde{t}_k \in A$ such that

$$\mathbf{M} \vDash \overline{\forall\!\!\!\exists}_{k+1} v_{k+1} \cdots \neg\sigma_{i,n}(v_0/t_0, \ldots, v_k/t_k, \ldots)\,.$$

On the other hand, if $\forall\!\!\!\exists_k$ is the quantifier "$\forall$", then

$$\widetilde{\mathsf{T}} \vdash \forall v_k \cdots \sigma_{i,n}(v_0/t_0, \ldots, v_k, \ldots)\,,$$

which implies, by L_{11},

$$\widetilde{\mathsf{T}} \vdash \forall\!\!\!\exists_{k+1} v_{k+1} \cdots \sigma_{i,n}(v_0/t_0, \dots, v_k/t_k, \dots) .$$

- If $\forall\!\!\!\exists_k$ is the quantifier "$\exists$", then, for $t_k \equiv (i, t_0, \dots, t_{k-1}, k)$,

$$\forall\!\!\!\exists_{k+1} v_{k+1} \cdots \sigma_{i,n}(v_0/t_0, \dots, v_k/t_k, \dots) \in \overline{\mathsf{T}}_c ,$$

which implies

$$\widetilde{\mathsf{T}} \vdash \forall\!\!\!\exists_{k+1} v_{k+1} \cdots \sigma_{i,n}(v_0/t_0, \dots, v_k/t_k, \dots) .$$

On the other hand, if $\forall\!\!\!\exists_k$ is the quantifier "$\exists$", then

$$\mathbf{M} \vDash \forall v_k \cdots \neg\sigma_{i,n}(v_0/t_0, \dots, v_k, \dots) ,$$

which implies, by L_{11},

$$\mathbf{M} \vDash \overline{\forall\!\!\!\exists}_{k+1} v_{k+1} \cdots \neg\sigma_{i,n}(v_0/t_0, \dots, v_k/t_k, \dots) .$$

Proceeding this way, we finally get

$$\mathbf{M} \vDash \neg\sigma_{i,n}(v_0/t_0, \dots, v_n/t_n) \quad \text{and} \quad \widetilde{\mathsf{T}} \vdash \sigma_{i,n}(v_0/t_0, \dots, v_n/t_n) .$$

Since the latter implies $\neg\sigma_{i,n}(v_0/t_0, \dots, v_n/t_n) \notin \widetilde{\mathsf{T}}$ and since $\sigma_{i,n}$ is variable-free, this is a contradiction to what we have proved above. $\dashv$

The following theorem just summarises what we have achieved so far:

COUNTABLE GÖDEL-HENKIN COMPLETENESS THEOREM 5.4. *If $\mathscr{L}$ is a countable signature and T is a consistent set of $\mathscr{L}$-sentences, then T has a model. Moreover, if $\mathsf{T} \nvdash \sigma_0$ (for some $\mathscr{L}$-sentence σ_0), then $\mathsf{T} + \neg\sigma_0$ has a model.*

In the next chapter, we shall prove the COMPLETENESS THEOREM for arbitrarily large signatures, but before, we would like to present a few consequences which follow directly from the COUNTABLE GÖDEL-HENKIN COMPLETENESS THEOREM (or its proof), or in combination with the COMPACTNESS THEOREM.

Some Consequences

Let $\mathscr{L}$ be a countable signature, T a set of $\mathscr{L}$-sentences, and σ_0 an $\mathscr{L}$-sentence.

- If $\mathsf{T} \nvdash \sigma_0$, then there is an $\mathscr{L}$-structure $\mathbf{M}$ such that $\mathbf{M} \vDash \mathsf{T} + \neg\sigma_0$:

$$\mathsf{T} \nvdash \sigma_0 \quad \Longrightarrow \quad \exists \mathbf{M} \left(\mathbf{M} \vDash \mathsf{T} + \neg\sigma_0 \right)$$

- If T is consistent, then T has a model:

$$\mathrm{Con}(\mathsf{T}) \quad \Longrightarrow \quad \exists \mathbf{M} \, (\mathbf{M} \vDash \mathsf{T})$$

- If each model of T is also a model of σ_0, then $\mathsf{T} \vdash \sigma_0$:

$$\forall \mathbf{M} \, (\mathbf{M} \vDash \mathsf{T} \Longrightarrow \mathbf{M} \vDash \sigma_0) \quad \Longrightarrow \quad \mathsf{T} \vdash \sigma_0$$

- In combination with the COMPACTNESS THEOREM 1.15 we get

$$\mathrm{Con}(\mathsf{T}) \quad \Longleftrightarrow \quad \exists \mathbf{M} \, (\mathbf{M} \vDash \mathsf{T})$$

  and finally:

$$\underbrace{\forall \mathbf{M} \, (\mathbf{M} \vDash \mathsf{T} \Longrightarrow \mathbf{M} \vDash \sigma_0)}_{\mathsf{T} \vDash \sigma_0} \quad \Longleftrightarrow \quad \mathsf{T} \vdash \sigma_0$$

The last consequence allows us to replace *formal proofs* with *mathematical proofs*: For example, instead of proving formally the uniqueness of the neutral element in groups from the axioms of Group Theory GT, we just show that in every model of GT (*i.e.*, in every group), the neutral element is unique. So, instead of $\mathsf{GT} \vdash \sigma_0$, we just show $\mathsf{GT} \vDash \sigma_0$.

As a last consequence we would like to mention the DOWNWARD LÖWENHEIM–SKOLEM THEOREM, which is also known as SKOLEM'S PARADOX.

DOWNWARD LÖWENHEIM-SKOLEM THEOREM 5.5. *If $\mathscr{L}$ is a countable signature and T is a consistent set of $\mathscr{L}$-sentences, then T has a countable model.*

Proof. In the previous chapter, we began with a countable signature $\mathscr{L}$ and a consistent set of $\mathscr{L}$-sentences T; and at the end, the domain A of the model $\mathbf{M}$ of T was a finite or potentially infinite list of lists. So, the model $\mathbf{M}$ we constructed is countable. $\dashv$

EXERCISES

8. Zeige mit Aufgabe 23.(b), dass die Theorie der dichten linearen Ordnungen vollständig ist; *d.h.* für alle $\mathscr{L}_{\mathsf{DLO}}$-Sätze σ gilt:

$$\textit{entweder} \quad \mathsf{DLO} \vdash \sigma \quad \textit{oder} \quad \mathsf{DLO} \vdash \neg\sigma$$

9. Prove FACT 5.2.

Chapter 6
The Completeness Theorem

Filters & Ultrafilters

Ultraproducts

Łoš's Theorem

The Gödel-Henkin Completeness Theorem

The Upward Löwenheim-Skolem Theorem

EXERCISES

10. eine Aufgabe

Chapter 7
Language Extensions by Definitions

Sometimes it is convenient to extend a given signature $\mathscr{L}$ by adding new non-logical symbols which have to be defined properly within the language $\mathscr{L}$ or with respect to a given $\mathscr{L}$-theory T. Let the extended signature be $\mathscr{L}^*$ and let the corresponding extended $\mathscr{L}^*$-theory be T^*. Since T is an $\mathscr{L}$-theory, we can just prove $\mathscr{L}$-sentences from T but no $\mathscr{L}^*$-sentences which contain symbols from $\mathscr{L}^*\backslash\mathscr{L}$. However, this does not imply that we can prove substantially more from T^* than from T: It might be that for each $\mathscr{L}^*$-sentence σ^* which is provable from T^* there is an $\mathscr{L}$-sentence $\widetilde{\sigma}$, such that $\mathsf{T}^* \vdash \sigma^* \leftrightarrow \widetilde{\sigma}$ and $\mathsf{T} \vdash \widetilde{\sigma}$; which is indeed the case as we shall see below.

Defining new Relation Symbols

Let us first consider an example from Peano Arithmetic: Extend the signature $\mathscr{L}_{\mathsf{PA}}$ of Peano Arithmetic by adding the binary relation symbol "$<$" and denote the extended signature by $\mathscr{L}^*_{\mathsf{PA}} := \mathscr{L}_{\mathsf{PA}} \cup \{<\}$. In order to define the binary relation "$<$", we give an $\mathscr{L}_{\mathsf{PA}}$-formula $\psi_<$ with two free variables (*e.g.*, x and y) and say that the relation $x < y$ holds if and only if $\psi_<(x,y)$ holds. In our case, $\psi_<(x,y) \equiv \exists z(x + \mathtt{s}z = y)$. So, we would define "$<$" by stipulating:

$$x < y :\Longleftrightarrow \exists z(x + \mathtt{s}z = y)$$

The problem is now to find for each $\mathscr{L}^*_{\mathsf{PA}}$-sentence σ^* an $\mathscr{L}_{\mathsf{PA}}$-sentence $\widetilde{\sigma}$ and an extension PA^* of PA, such that $\mathsf{PA}^* \vdash \sigma^* \leftrightarrow \widetilde{\sigma}$ and whenever $\mathsf{PA}^* \vdash \sigma^*$, then $\mathsf{PA} \vdash \widetilde{\sigma}$.

The following result provides an algorithm which transforms sentences σ^* in the extended language into equivalent sentences $\widetilde{\sigma}$ in the original language:

THEOREM 7.1. *Let $\mathscr{L}$ be a signature, let R be an n-ary relation symbol which does not belong to $\mathscr{L}$, and let $\mathscr{L}^* := \mathscr{L} \cup \{R\}$. Furthermore, let $\psi_R(v_1,\dots,v_n)$ be an $\mathscr{L}$-formula with* $\mathrm{free}(\psi_R) = \{v_1,\dots,v_n\}$ *and let*

$$\vartheta_R \equiv \forall v_1 \cdots \forall v_n \big(Rv_1 \cdots v_n \leftrightarrow \psi_R(v_1, \ldots, v_n)\big)\,.$$

Finally, let T *be a consistent* $\mathscr{L}$*-theory and let* $\mathsf{T}^* := \mathsf{T} + \vartheta_R$.

Then there exists an effective algorithm which transforms each $\mathscr{L}^*$*-formula* φ^* *into an* $\mathscr{L}$*-formula* $\widetilde{\varphi}$*, such that:*

(a) *If* R *does not appear in* φ^**, then* $\widetilde{\varphi} \equiv \varphi^*$.

(b) $\widetilde{\neg\varphi} \equiv \neg\widetilde{\varphi}$ *(for* $\varphi^* \equiv \neg\varphi$*)*

(c) $\widetilde{\wedge\varphi_1\varphi_2} \equiv \wedge\widetilde{\varphi}_1\widetilde{\varphi}_2$ *(for* $\varphi^* \equiv \wedge\varphi_1\varphi_2$*)*

(d) $\widetilde{\exists\nu\varphi} \equiv \exists\nu\widetilde{\varphi}$ *(for* $\varphi^* \equiv \exists\nu\varphi$*)*

(e) $\mathsf{T}^* \vdash \varphi^* \leftrightarrow \widetilde{\varphi}$

(f) *If* $\mathsf{T}^* \vdash \varphi^*$*, then* $\mathsf{T} \vdash \widetilde{\varphi}$.

Proof. Let φ^* be an arbitrary $\mathscr{L}^*$-formula. In φ^* we replace each occurrence of $R\nu_1 \cdots \nu_n$ with a formula $\psi'_R(\nu_1, \ldots, \nu_n)$ such that

$$\psi'_R(\nu_1, \ldots, \nu_n) \Leftrightarrow \psi_R(\nu_1, \ldots, \nu_n)$$

and no variable $\nu_1, \ldots, \nu_n$ is bounded in ψ'_R. For the resulting $\mathscr{L}$-formula $\widetilde{\varphi}$, (a)–(d) are obviously satisfied and it remains to prove (e) & (f).

To prove (e), by the GÖDEL-HENKIN COMPLETENESS THEOREM it is enough to show that $\varphi^* \leftrightarrow \widetilde{\varphi}$ holds in every model $\mathbf{M}^*$ of T^*. So, let $\mathbf{M}^*$ be an arbitrary model of T^*. In particular, $\mathbf{M}^* \vDash \vartheta_R$. If φ^* does not contain R, then we are done. Otherwise, if φ^* is atomic, then $\varphi^* \equiv Rt_1 \cdots t_n$ for some $\mathscr{L}$-terms $t_1, \ldots, t_n$. Since $\mathbf{M}^* \vDash \vartheta_R$, we get

$$\mathbf{M}^* \vDash Rt_1 \cdots t_n \leftrightarrow \psi'_R(t_1, \ldots, t_n)\,.$$

This shows $\mathbf{M}^* \vDash \varphi^* \leftrightarrow \widetilde{\varphi}$ for atomic formulas and by (b)–(d) we get the result for arbitrary formulas.

For (f), we first notice that every model $\mathbf{M}$ of T can be extended to an $\mathscr{L}^*$-structure $\mathbf{M}^*$ such that $\mathbf{M}^* \vDash \mathsf{T}^*$. Let $\mathbf{M}$ be an abitrary model of T and let $\mathbf{M}^*$ be such an extension to a model of T^*. By (e), for each $\mathscr{L}^*$-formula φ^* we have

$$\mathbf{M}^* \vDash \varphi^* \quad \Longleftrightarrow \quad \mathbf{M}^* \vDash \widetilde{\varphi}\,.$$

Now, if $\mathsf{T}^* \vdash \varphi^*$, then $\mathbf{M}^* \vDash \varphi^*$, which implies that $\mathbf{M}^* \vDash \widetilde{\varphi}$. Since $\widetilde{\varphi}$ is an $\mathscr{L}$-formula, we get $\mathbf{M} \vDash \widetilde{\varphi}$, and since the model $\mathbf{M}$ of T was arbitrary, by the GÖDEL-HENKIN COMPLETENESS THEOREM we get $\mathsf{T} \vdash \widetilde{\varphi}$. $\dashv$

Defining new Function Symbols

The situation is slightly more subtle if we define new functions. However, there is also an algorithm which transforms sentences σ^* in the extended language into equivalent sentences $\widetilde{\sigma}$ in the original language:

THEOREM 7.2. *Let $\mathscr{L}$ be a signature, let f be an n-ary relation symbol which does not belong to $\mathscr{L}$, let $\mathscr{L}^* := \mathscr{L} \cup \{f\}$ and let T be a consistent $\mathscr{L}$-theory. Furthermore, let $\psi_f(v_1, \ldots, v_n, y)$ be an $\mathscr{L}$-formula with $\mathrm{free}(\psi_f) = \{v_1, \ldots, v_n, y\}$ such that*

$$\mathsf{T} \vdash \forall v_1 \cdots \forall v_n \exists! y\, \psi_f(v_1, \ldots, v_n, y)\,.$$

Finally, let

$$\vartheta_f \equiv \forall v_1 \cdots \forall v_n \forall y \big(f v_1 \cdots v_n = y \leftrightarrow \psi_f(v_1, \ldots, v_n, y) \big)$$

and let $\mathsf{T}^* := \mathsf{T} + \vartheta_f$.

Then there exists an effective algorithm which transforms each $\mathscr{L}^*$-formula φ^* into an $\mathscr{L}$-formula $\widetilde{\varphi}$, such that:

(a) *If f does not appear in φ^*, then $\widetilde{\varphi} \equiv \varphi^*$.*

(b) $\widetilde{\neg\varphi} \equiv \neg\widetilde{\varphi}$ *(for $\varphi^* \equiv \neg\varphi$)*

(c) $\widetilde{\wedge\varphi_1\varphi_2} \equiv \wedge\widetilde{\varphi}_1\widetilde{\varphi}_2$ *(for $\varphi^* \equiv \wedge\varphi_1\varphi_2$)*

(d) $\widetilde{\exists\nu\varphi} \equiv \exists\nu\widetilde{\varphi}$ *(for $\varphi^* \equiv \exists\nu\varphi$)*

(e) $\mathsf{T}^* \vdash \varphi^* \leftrightarrow \widetilde{\varphi}$

(f) *If $\mathsf{T}^* \vdash \varphi^*$, then $\mathsf{T} \vdash \widetilde{\varphi}$.*

Proof. By an *elementary f-term* we mean an $\mathscr{L}^*$-term of the form $f t_1 \cdots t_n$, where $t_1, \ldots, t_n$ are $\mathscr{L}^*$-terms which do not contain the symbol f. We first prove the theorem for atomic $\mathscr{L}^*$-formulae φ^* (*i.e.*, for formulae which are free of quantifiers and logical operators). Let $\varphi^*(f|w)$ be the result of replacing the leftmost occurence of an elementary f-term in φ^* with a new symbol w, which stands for a new variable. Then, the formula

$$\exists w \big(\psi_f(t_1, \ldots, t_n, w) \wedge \varphi^*(f|w) \big)$$

is called the *f-transform of* φ^*. If φ^* does not contain f, then let φ^* be its own f-tranform. Before we procceed, let us prove the following

CLAIM. $\mathsf{T}^* \vdash \exists w \big(\psi_f(t_1, \ldots, t_n, w) \wedge \varphi^*(f|w) \big) \leftrightarrow \varphi^*$

Proof of Claim. Let $\mathbf{M}^*$ be a model of T^* with domain A, let j be an arbitrary assignment which assigns to w an element of A, and let $\mathbf{M}^*_j := (\mathbf{M}^*, j)$ be the corresponding $\mathscr{L}^*$-interpretation.

Assume that
$$\mathbf{M}^*_j \vDash \exists w\big(\psi_f(t_1,\dots,t_n,w) \wedge \varphi^*(f|w)\big)\,.$$
Then, since $\mathsf{T}^* \vdash \forall v_1 \cdots \forall v_n \exists! y\, \psi_f(v_1,\dots,v_n,y)$, there exists a unique $b \in A$ such that
$$\mathbf{M}^*_{j\frac{b}{w}} \vDash \psi_f(t_1,\dots,t_n,w) \wedge \varphi^*(f|w)\,,$$
which is the same as saying that
$$\mathbf{M}^*_j \vDash \psi_f(t_1,\dots,t_n,b) \wedge \varphi^*(f|b)\,.$$
Now, since $\mathbf{M}^*_j \vDash \vartheta_f$, b is the same object as $f^{\mathbf{M}^*_j} t_1^{\mathbf{M}^*_j} \cdots t_n^{\mathbf{M}^*_j}$. This implies
$$\mathbf{M}^*_j \vDash f t_1 \cdots t_n = b\,,$$
and shows that
$$\mathbf{M}^*_j \vDash \varphi^*\,.$$
For the reverse implication assume that $\mathbf{M}^*_j \vDash \varphi^*$ and let b be the same object as $f^{\mathbf{M}^*_j} t_1^{\mathbf{M}^*_j} \cdots t_n^{\mathbf{M}^*_j}$. Then $\mathbf{M}^*_j \vDash \varphi^*(f|b)$ and, since $\mathbf{M}^*_j \vDash \vartheta_f$,
$$\mathbf{M}^*_j \vDash \psi_f(t_1,\dots,t_n,w) \leftrightarrow f t_1 \cdots t_n = w\,.$$
In particular we get
$$\mathbf{M}^*_{j\frac{b}{w}} \vDash \psi_f(t_1,\dots,t_n,b) \leftrightarrow f t_1 \cdots t_n = b\,,$$
and because $f^{\mathbf{M}^*_j} t_1^{\mathbf{M}^*_j} \cdots t_n^{\mathbf{M}^*_j}$ is the same object as b, we get $\mathbf{M}^*_j \vDash \psi_f(t_1,\dots,t_n,b)$, and since we already know $\mathbf{M}^*_j \vDash \varphi^*(f|b)$, we have
$$\mathbf{M}^*_j \vDash \psi_f(t_1,\dots,t_n,b) \wedge \varphi^*(f|b)\,.$$
So, there exists a b in A, such that
$$\mathbf{M}^*_{j\frac{b}{w}} \vDash \psi_f(t_1,\dots,t_n,w) \wedge \varphi^*(f|w)\,,$$
which is the same as saying that
$$\mathbf{M}^*_j \vDash \exists w\big(\psi_f(t_1,\dots,t_n,w) \wedge \varphi^*(f|w)\big)\,.$$
Since the model $\mathbf{M}^*$ of T^* was arbitrary, by the GÖDEL-HENKIN COMPLETENESS THEOREM we get $\mathsf{T}^* \vdash \exists w\big(\psi_f(t_1,\dots,t_n,w) \wedge \varphi^*(f|w)\big) \leftrightarrow \varphi^*$. $\dashv_{\text{Claim}}$

Since the f-transform $\exists w\big(\psi_f(t_1,\dots,t_n,w) \wedge \varphi^*(f|w)\big)$ of φ^* contains one less f than φ^*, if we take successive f-transforms (introducing always new variables), eventually we obtain an an atomic $\mathscr{L}$-formula $\widetilde{\varphi}$ (*i.e.*, a formula which does not contain f) such that $\mathsf{T}^* \vdash \varphi^* \leftrightarrow \widetilde{\varphi}$. We call $\widetilde{\varphi}$ the *f-less transform of* φ^*.

In order to get f-less transforms of non-atomic $\mathscr{L}^*$-formulae φ^*, we just extend the definition by letting $\widetilde{\neg\varphi}$ be $\neg\widetilde{\varphi}$, $\widetilde{\wedge\varphi_1\varphi_2}$ be $\wedge\widetilde{\varphi}_1\widetilde{\varphi}_2$, and $\widetilde{\exists\nu\varphi}$ be $\exists\nu\widetilde{\varphi}$; properties (a)–(e) are then obvious.

It remains to prove property (f). Let $\mathbf{M}_0$ be an abitrary model of T with domain A. Then, since $\mathsf{T} \vdash \forall v_1 \cdots \forall v_n \exists! y\, \psi_f(v_1, \ldots, v_n, y)$, for all $a_1, \ldots, a_n$ in A there exists a unique b in A such that

$$\mathbf{M}_0 \vDash \psi_f(a_1, \ldots, a_n, b)$$

and we define the n-ary function f^* on A by stipulating:

$$f^*(a_1, \ldots, a_n) := b$$

With this definition, we can extend the $\mathscr{L}$-structure $\mathbf{M}_0$ to an $\mathscr{L}^*$-structure $\mathbf{M}_0^*$, where we still have $\mathbf{M}^* \vDash \mathsf{T}$. With the definition of f^* we get in addition $\mathbf{M}_0^* \vDash \vartheta_f$, which implies $\mathbf{M}_0^* \vDash \mathsf{T}^*$. If we have $\mathsf{T}^* \vdash \varphi^*$, for some $\mathscr{L}^*$-formula φ^*, then there exists an $\mathscr{L}$-formula $\widetilde{\varphi}$, such that $\mathsf{T}^* \vdash \varphi^* \leftrightarrow \widetilde{\varphi}$, *i.e.*, $\mathsf{T}^* \vdash \widetilde{\varphi}$. Since $\mathsf{T}^* \vdash \widetilde{\varphi}$ implies $\mathbf{M}_0^* \vDash \widetilde{\varphi}$, and because $\widetilde{\varphi}$ is an $\mathscr{L}$-formula, we have $\mathbf{M}_0 \vDash \widetilde{\varphi}$. Now, since the model $\mathbf{M}_0$ of T was arbitrary, by the GÖDEL-HENKIN COMPLETENESS THEOREM we get $\mathsf{T} \vdash \widetilde{\varphi}$. $\dashv$

Defining new Constant Symbols

Constant symbols can be handled like 0-are function symbols:

FACT 7.3. *Let $\mathscr{L}$ be a signature, let c be constant symbol which does not belong to $\mathscr{L}$, let $\mathscr{L}^* := \mathscr{L} \cup \{c\}$ and let T be a consistent $\mathscr{L}$-theory. Furthermore, let $\psi_c(y)$ be an $\mathscr{L}$-formula with $\mathrm{free}(\psi_c) = \{y\}$ such that*

$$\mathsf{T} \vdash \exists! y\, \psi_c(y)\,.$$

Finally, let

$$\vartheta_c \equiv \forall y\big(c = y \leftrightarrow \psi_c(y)\big)$$

and let $\mathsf{T}^* := \mathsf{T} + \vartheta_c$.

Then there exists an effective algorithm which transforms each $\mathscr{L}^*$-formula φ^* into an $\mathscr{L}$-formula $\widetilde{\varphi}$, such that:

(a) *If f does not appear in φ^*, then $\widetilde{\varphi} \equiv \varphi^*$.*

(b) $\widetilde{\neg\varphi} \equiv \neg\widetilde{\varphi}$ *(for $\varphi^* \equiv \neg\varphi$)*

(c) $\widetilde{\wedge\varphi_1\varphi_2} \equiv \wedge\widetilde{\varphi}_1\widetilde{\varphi}_2$ *(for $\varphi^* \equiv \wedge\varphi_1\varphi_2$)*

(d) $\widetilde{\exists\nu\varphi} \equiv \exists\nu\widetilde{\varphi}$ *(for $\varphi^* \equiv \exists\nu\varphi$)*

(e) $\mathsf{T}^* \vdash \varphi^* \leftrightarrow \widetilde{\varphi}$

(f) *If* $\mathsf{T}^* \vdash \varphi^*$, *then* $\mathsf{T} \vdash \widetilde{\varphi}$.

Proof. The algorithm is constructed in exactly the same way as in the proof of THEOREM 7.2. $\dashv$

EXERCISES

11. Something else

Part III
Gödel's Incompleteness Theorems

On the syntactical level, an $\mathscr{L}$-theory T is complete if for every $\mathscr{L}$-sentence σ, either $\mathsf{T} \vdash \sigma$ or $\mathsf{T} \vdash \neg\sigma$. On the semantical level, a consistent $\mathscr{L}$-theory is T complete if any two models of T are elementary equivalent.

In this part of the book we shall first provide a few models of Peano Arithmetic PA, where we assume that PA is consistent. Then, we shall prove GÖDELS FIRST INCOMPLETENESS THEOREM, which states that Peano Arithmetic PA is not complete, *i.e.*, there is a $\mathscr{L}_{\mathsf{PA}}$-sentence σ, such that neither $\mathsf{PA} \vdash \sigma$ nor $\mathsf{PA} \vdash \neg\sigma$. In a second step we shall prove GÖDELS SECOND INCOMPLETENESS THEOREM which shows that the statement $\mathrm{Con}(\mathsf{PA})$ can be formalised in PA, but it cannot be proved in PA (unless PA is inconsistent).

Chapter 8
Models of Peano Arithmetic

By the COMPLETENESS THEOREM we know that every consistent theory T has a model, and if T has an infinite model, then it has also arbitrarily large models. So, if we assume that Peano Arithmetic PA is consistent—what seems sensible—then there exists a model of PA, and because this model is infinite, PA must have arbitrarily large models as well.

In this chapter we provide a few models of PA. We shall begin by constructing the so-called *standard model*, then we shall extend this model to a countable *non-standard model*, and finally we shall construct uncountable models of PA.

The Standard Model of Peano Arithmetic

For the sake of completeness, let us first recall the language and the seven axioms of Peano Arithmetic PA:

PA: The language PA is $\mathscr{L}_{\mathsf{PA}} = \{0, \mathtt{s}, +, \cdot\}$, where "0" is a constant symbol, "$\mathtt{s}$" is a unary function symbol, and "$+$" & "$\cdot$" are binary function symbols.

PA_0: $\neg\exists x(\mathtt{s}x = 0)$
PA_1: $\forall x\forall y(\mathtt{s}x = \mathtt{s}y \to x = y)$,
PA_2: $\forall x(x + 0 = x)$
PA_3: $\forall x\forall y(x + \mathtt{s}y = \mathtt{s}(x + y))$
PA_4: $\forall x(x \cdot 0 = 0)$
PA_5: $\forall x\forall y(x \cdot \mathtt{s}y = (x \cdot y) + x)$

If φ is any $\mathscr{L}_{\mathsf{PA}}$-formula with $x \in \mathrm{free}(\varphi)$, then:

PA_6: $\big(\varphi(0) \wedge \forall x(\varphi(x) \to \varphi(\mathtt{s}(x)))\big) \to \forall x\varphi(x)$

The domain $\mathbb{N}$ of our standard model is essentially the same as the list $\mathbb{N}$ introduced in Chapter 0, just the elements are named differently. The objects in $\mathbb{N}$ are defined as strings of symbols resulting from applying F I N I T E L Y many times the following rules:

(N0) 0 is a string.
(N1) If we have already built the string ξ, then $\mathsf{s}\xi$ is also a string.

For each string $\xi \in \mathbb{N}$ we have:

$$\text{either} \quad \xi \equiv 0 \quad \text{or} \quad \xi \equiv \underbrace{\mathsf{s} \cdots \mathsf{s}}_{\substack{\text{non-empty}\\ \text{finite string}}} 0$$

To each non-empty finite string which consists just of the symbol s we assign a kind of "length" m and write $\underset{m}{\mathsf{s}}$ instead of $\mathsf{s} \cdots \mathsf{s}$. So, $\underset{m}{\mathsf{s}}$ is just an abbreviation of a finite string of the form $\mathsf{s} \cdots \mathsf{s}$.

REMARK. With this notation we get that each string in $\mathbb{N}$ is either 0 or of the form $\underset{m}{\mathsf{s}}0$. Further we get that for any strings $\underset{m}{\mathsf{s}}$ and $\underset{n}{\mathsf{s}}$ we have for example

$$\underset{m}{\mathsf{s}}\underset{n}{\mathsf{s}}0 \equiv \underset{n}{\mathsf{s}}\underset{m}{\mathsf{s}}0\,, \qquad \mathsf{s}\underset{m}{\mathsf{s}}\underset{n}{\mathsf{s}}0 \equiv \mathsf{s}\underset{n}{\mathsf{s}}\underset{m}{\mathsf{s}}0\,, \qquad \mathsf{s}\underset{m}{\mathsf{s}}\underset{n}{\mathsf{s}}0 \equiv \underset{m}{\mathsf{s}}\mathsf{s}\underset{n}{\mathsf{s}}0\,,$$

and further we get:

$$\underset{m}{\mathsf{s}}0 \equiv \underset{n}{\mathsf{s}}0 \iff \mathsf{s}\underset{m}{\mathsf{s}}0 \equiv \mathsf{s}\underset{n}{\mathsf{s}}0$$
$$\underset{m}{\mathsf{s}}0 \equiv \underset{n}{\mathsf{s}}0 \iff \underset{n}{\mathsf{s}}\underset{m}{\mathsf{s}}0 \equiv \underset{m}{\mathsf{s}}\underset{n}{\mathsf{s}}0$$

This we can deduce from Euclid: EUKLID erwähnen, Buch I, Axiom 2

Now, we are going to define the standard model of PA with domain $\mathbb{N}$. For this, we have to define first an $\mathscr{L}_{\mathsf{PA}}$-structure $\mathbb{N}$, which means, that we have to interpret the non-logical symbols in $\mathscr{L}_{\mathsf{PA}}$:

$$0^{\mathbb{N}} := 0$$

$$\begin{aligned} \mathsf{s}^{\mathbb{N}} :\ \mathbb{N} &\to \mathbb{N} \\ 0 &\mapsto \mathsf{s}0 \\ \underset{m}{\mathsf{s}}0 &\mapsto \mathsf{s}\underset{m}{\mathsf{s}}0 \end{aligned}$$

$$\begin{aligned} +^{\mathbb{N}} :\quad \mathbb{N} \times \mathbb{N} &\to \mathbb{N} \\ \langle 0, 0 \rangle &\mapsto 0 \\ \langle \underset{m}{\mathsf{s}}0, 0 \rangle &\mapsto \underset{m}{\mathsf{s}}0 \\ \langle 0, \underset{n}{\mathsf{s}}0 \rangle &\mapsto \underset{n}{\mathsf{s}}0 \\ \langle \underset{m}{\mathsf{s}}0, \underset{n}{\mathsf{s}}0 \rangle &\mapsto \underset{m}{\mathsf{s}}\underset{n}{\mathsf{s}}0 \end{aligned}$$

$$
\begin{aligned}
\cdot^{\mathbb{N}} : \quad \mathbb{N} \times \mathbb{N} &\to \mathbb{N} \\
\langle 0, 0 \rangle &\mapsto 0 \\
\langle \underset{m}{\mathtt{s}}0, 0 \rangle &\mapsto 0 \\
\langle 0, \underset{n}{\mathtt{s}}0 \rangle &\mapsto 0 \\
\langle \underset{m}{\mathtt{s}}0, \underset{n}{\mathtt{s}}0 \rangle &\mapsto \underbrace{\begin{matrix} \underset{n}{\mathtt{s}}\,\underset{n}{\mathtt{s}} & \cdots & \underset{n}{\mathtt{s}}0 \\ \uparrow\uparrow & \cdots & \uparrow \\ \mathtt{s}\,\mathtt{s} & \cdots & \mathtt{s} \end{matrix}}_{\underset{m}{\mathtt{s}}}
\end{aligned}
$$

The main feature of the $\mathscr{L}_{\mathsf{PA}}$-structure $\mathbb{N}$ is that every element of $\mathbb{N}$ corresponds to a certain $\mathscr{L}_{\mathsf{PA}}$-term. In order to prove this, we introduce the following notion: Like for elements of $\mathbb{N}$, to each non-empty finite string $\mathtt{s} \cdots \mathtt{s}$, which consists just of the unary function symbol $\mathtt{s} \in \mathscr{L}_{\mathsf{PA}}$, we assign again a kind of "length" m and write $\underset{m}{\mathtt{s}}$ instead of $\mathtt{s} \cdots \mathtt{s}$. So, $\underset{m}{\mathtt{s}}$ is just an abbreviation of a finite string of the form $\mathtt{s} \cdots \mathtt{s}$. As a consequence of this definition we get the following

FACT 8.1. *For all finite strings $\underset{m}{\mathtt{s}}$ and $\underset{n}{\mathtt{s}}$ we have:*

(a) $\mathsf{PA} \vdash \underset{m}{\mathtt{s}}\mathtt{0} \neq \mathtt{0}$

(b) $\mathsf{PA} \vdash \underset{m}{\mathtt{s}}\mathtt{0} = \underset{n}{\mathtt{s}}\mathtt{0} \iff \underset{m}{\mathtt{s}}\mathtt{0} \equiv \underset{n}{\mathtt{s}}\mathtt{0}$

Proof. (a) follows from PA_0, and (b) follows from PA_1 and L_{15}. ⊣

LEMMA 8.2. *Every element of $\mathbb{N}$ corresponds to a unique finite application of the function $\mathtt{s}$ to $\mathtt{0}$. More formally, for every element $\underset{m}{\mathtt{s}}0$ of $\mathbb{N}$ there is a unique $\mathscr{L}_{\mathsf{PA}}$-term $\underset{m}{\mathtt{s}}\mathtt{0}$ such that*

$$\underset{m}{\mathtt{s}}\mathtt{0}^{\mathbb{N}} \quad \text{IS THE SAME OBJECT AS} \quad \underset{m}{\mathtt{s}}0\,.$$

Proof. By definition of $\mathtt{s}^{\mathbb{N}}$ we get that $\mathtt{s}^{\mathbb{N}}(\underset{m}{\mathtt{s}}0)$ is the same element of $\mathbb{N}$ as $\mathtt{s}\underset{m}{\mathtt{s}}0$, and after applying this fact finitely many times we get:

$$
\overbrace{\underbrace{\begin{matrix} \mathtt{s}^{\mathbb{N}} & \mathtt{s}^{\mathbb{N}} & \cdots & \mathtt{s}^{\mathbb{N}} & \mathtt{0}^{\mathbb{N}} \\ \updownarrow & \updownarrow & \cdots & \updownarrow & \updownarrow \\ \mathtt{s} & \mathtt{s} & \cdots & \mathtt{s} & 0 \end{matrix}}_{\underset{m}{\mathtt{s}}0}}^{\underset{m}{\mathtt{s}}\mathtt{0}^{\mathbb{N}}}
$$

The uniqueness of $\underset{m}{\mathtt{s}}\mathtt{0}$ follows from FACT 8.1. ⊣

Now, we are ready to prove that the $\mathscr{L}_{\mathsf{PA}}$-structure $\mathbb{N}$, which is called the **standard model** of Peano Arithmetic, is indeed a model of PA.

THEOREM 8.3. $\mathbb{N} \vDash \mathsf{PA}$

Proof. By definition of $\mathtt{s}^{\mathbb{N}}$ we get $\mathbb{N} \vDash \mathsf{PA}_0$ and by the REMARK above we also have $\mathbb{N} \vDash \mathsf{PA}_1$. Further, by definition of $+^{\mathbb{N}}$ and $\cdot^{\mathbb{N}}$ we get $\mathbb{N} \vDash \mathsf{PA}_2$ and $\mathbb{N} \vDash \mathsf{PA}_4$ respectively. For PA_3 notice first that by definition of $+^{\mathbb{N}}$ we get $0 +^{\mathbb{N}} \mathtt{s}\underset{n}{\mathtt{s}}0 \equiv \mathtt{s}\underset{n}{\mathtt{s}}0$, where $\mathtt{s}\underset{n}{\mathtt{s}}0 \equiv \mathtt{s}^{\mathbb{N}}(0 +^{\mathbb{N}} \underset{n}{\mathtt{s}}0)$. Let us now compute $\underset{m}{\mathtt{s}}0 +^{\mathbb{N}} \mathtt{s}\underset{n}{\mathtt{s}}0$: By definition of $+^{\mathbb{N}}$ and the REMARK above we get

$$\underset{m}{\mathtt{s}}0 +^{\mathbb{N}} \mathtt{s}\underset{n}{\mathtt{s}}0 \;\equiv\; \underset{m}{\mathtt{s}}\mathtt{s}\underset{n}{\mathtt{s}}0 \quad \text{and} \quad \underset{m}{\mathtt{s}}\mathtt{s}\underset{n}{\mathtt{s}}0 \;\equiv\; \mathtt{s}\underset{m\,n}{\mathtt{s}}0\,,$$

where

$$\mathtt{s}\underset{m\,n}{\mathtt{s}}0 \;\equiv\; \mathtt{s}^{\mathbb{N}}(\underset{m}{\mathtt{s}}0 + \underset{n}{\mathtt{s}}0)\,.$$

Similarly, we can prove $\mathbb{N} \vDash \mathsf{PA}_5$ (see EXERCISE 12). In order to show that $\mathbb{N} \vDash \mathsf{PA}_6$, let $\varphi(x)$ be an $\mathscr{L}_{\mathsf{PA}}$-formula and let us assume that

$$\mathbb{N} \vDash \varphi(\mathtt{0}) \wedge \forall x\big(\varphi(x) \to \varphi(\mathtt{s}x)\big)\,.$$

We have to show that $\mathbb{N} \vDash \forall x \varphi(x)$. By definition of models we get that $\varphi(0)$ holds in $\mathbb{N}$ and for all $\xi \in \mathbb{N}$: if $\varphi(\xi)$ holds in $\mathbb{N}$, then also $\varphi(\mathtt{s}^{\mathbb{N}}\xi)$ holds in $\mathbb{N}$. Now, let $\underset{m}{\mathtt{s}}0$ be an arbitrary element of $\mathbb{N}$. Then, by LEMMA 8.2, we know that $\underset{m}{\mathtt{s}}0$ is the same object as $\underset{m}{\mathtt{s}}\mathtt{0}^{\mathbb{N}}$, which implies that $\varphi(\underset{m}{\mathtt{s}}0)$ holds in $\mathbb{N}$. Finally, since $\underset{m}{\mathtt{s}}0$ was arbitrary, we get that for *all* $\xi \in \mathbb{N}$, $\varphi(\xi)$ holds in $\mathbb{N}$; hence, $\mathbb{N} \vDash \forall x \varphi(x)$. $\dashv$

As a matter of fact we would like to mention that every model of PA must contain an isomorphic copy of the standard model $\mathbb{N}$. So, it would also make sense to call $\mathbb{N}$ the **minimal model** of of Peano Arithmetic. However, it is not clear so far, whether $\mathbb{N}$ is, up to isomorphisms, the only model of PA. This is not the case, as we shall see now.

Countable Non-Standard Models of Peano Arithmetic

Non-standard models which are elementarily equivalent to the standard model.

Uncountable Models of Peano Arithmetic

EXERCISES

12. Prove that $\mathbb{N} \vDash \mathsf{PA}_5$.