## First Order Logic or Predicate Calculus

Compared to Propositional Calculus, in First Order Logic we relieve to some extent the requirement to abstract from the content of particular statements and arguments and to concentrate upon their form, only. Namely, we admit that all the statements and arguments deal with some *objects*, that these objects have some *properties* or enter some *relations*, that some objects are explicitly mentioned by their *names* and, finally, that from given objects some new objects can be produced by means of certain *operations*. Though this shift of focus makes sense equally for natural languages as well as for languages of scientific theories, we will restrict our attention to the languages of *mathematical theories* exclusively, where such an approach is rather natural and its fruitfulness can be demonstrated in a convincing way. On the other hand, the reduction of the logical structure of natural languages as well as of the languages of most scientific theories to its fragment fitting within the framework set by the First Order Logic would turn out rather artificial.

First Order Logic, also called (Lower) Predicate Calculus, examines the structure of arguments and proofs used in mathematics, more precisely in mathematical theories describing classes of mathematical structures formed by sets of objects endowed with various finitary operations and relations, singled out by certain axioms. That way First Order Logic is mathematical logic both by its methods as well as by its subject.

We intend to develop and present the First Order Logic in a way parallel, as much as possible, to our previous development and presentation of Propositional Calculus. Consequently, the reader should see clearly both the similarities as well as the differences between these two branches of mathematical logic.

## First Order Languages and First Order Structures

A typical mathematical structure consists of a nonempty base set $A$ of objects, equipped with some finitary operations, some distinguished elements and some finitary relations.

**Example.** Number systems with operations of addition $+$, multiplication $\cdot$, distinguished elements 0 and 1 and (with the exception of complex numbers) the ordering relation $<$ form mathematical structures commonly denoted as follows:

Natural numbers $(\mathbb{N};\, +, \cdot, 0, 1, <)$
Integers $(\mathbb{Z};\, +, \cdot, 0, 1, <)$
Rational numbers $(\mathbb{Q};\, +, \cdot, 0, 1, <)$
Real numbers $(\mathbb{R};\, +, \cdot, 0, 1, <)$
Complex numbers $(\mathbb{C};\, +, \cdot, 0, 1)$

A *first order language* $L = (F, C, R, \nu)$ is given by some (maybe void) sets $F$, $C$, $R$ of functional (operation) symbols, constant symbols and relational (predicate) symbols, respectively, together with an arity function (signature) $\nu\colon F \cup R \to \mathbb{N}$, assigning to any symbol $s \in F \cup R$ its *arity* $\nu(s) \geq 1$. These are the *specific symbols* of $L$. Constant symbols are sometimes considered as nullary functional symbols, i.e., as elements $f$ of the set $F$ subject to $\nu(f) = 0$ (in that case the set $C$ does not explicitly occur in the description of $L$).

All first order languages contain common *logical symbols*:

- Object variables (or just variables): $x$, $y$, $z$, $u$, $v$, $w$, $x_0$, $x_1$, $x_2$, $\dots$ , $y'$, $y''$, $\dots$
- Logical connectives: $\neg$ (*not*), $\wedge$ (*and*), $\vee$ (*or*), $\Rightarrow$ (*if* $\dots$ *then* or *implies*), $\Leftrightarrow$ (*if and only if*) (two would suffice)
- Quantifiers: $\forall$ (*universal quantifier*), $\exists$ (*existential quantifier*) (one would suffice)
- Equality sign: $=$
- Auxiliary symbols: ( , ) (*parentheses*), , (*comma*) (they could be avoided)

**Remark.** At a glance it could seem that we should require the sets $F$, $C$, $R$ to be at most countable, since it makes no sense to admit that the language $L$ contains uncountably many specific symbols. Such a restriction, however, would bring us no technical advantage. More important, as we shall see later on, e.g., when dealing with various *diagrams* of structures, the methods and results of the study of uncountable languages have applications even for structures of countable first order languages.

A *first order structure*, i.e., a structure of some first order language $L = (F, C, R, \nu)$, briefly, an $L$-structure, $\mathcal{A} = (A; I)$ consists of a nonempty set $A$ (base set or carrier) and an interpretation $I$ of the specific symbols of $L$ in $A$:

- for $f \in F$, such that $\nu(f) = n$, $f^I \colon A^n \to A$
  (each $n$-ary operation symbol is interpreted as an $n$-ary operation on $A$)
- for $c \in C$, $c^I \in A$
  (each constant symbol is interpreted as some distinguished element of $A$)
- for $r \in R$, such that $\nu(r) = n$, $r^I \subseteq A^n$
  (each $n$-ary relation symbol is interpreted as an $n$-ary relation on $A$)

Instead of $s^I$ we frequently write $s^{\mathcal{A}}$ or just $s$ for any specific symbol $s$.

## Terms and Formulas

*Terms* of a first order language $L$, briefly $L$-terms, are composed of variables, constant symbols and functional symbols of $L$. The set $\mathrm{Term}(L)$ of all $L$-terms is the smallest set such that

1° $x \in \mathrm{Term}(L)$ for each variable $x$ (every variable is a term);

2° $c \in \mathrm{Term}(L)$ for each $c \in C$ (every constant symbol is a term);

3° if $f \in F$, $\nu(f) = n$, and $t_1, \dots, t_n \in \mathrm{Term}(L)$, then $f(t_1, \dots, t_n) \in \mathrm{Term}(L)$ (if $f$ is an $n$-ary functional symbol and $t_1, \dots, t_n$ are terms, then $f(t_1, \dots, t_n)$ is a term, as well).

This is a *recursive definition.* Terms in 1°, 2°, i.e., variables and constant symbols, are called *atomic terms.* In order to prove that some set of $L$-terms contains all the $L$-terms, it suffices to show that it fulfills all the three conditions 1°, 2° and 3°.

For a binary operation symbol $f$ we usually write $t_1 f t_2$ instead of $f(t_1, t_2)$ in 3°.

*Constant terms* are terms containing no variables. If all the variables occurring in a term $t$ are contained in the list $x_1, \dots, x_k$, we write $t(x_1, \dots, x_k)$. There is one exception: writing $t$ doesn't necessarily mean that $t$ is a constant term.

Given an $L$-structure $\mathcal{A} = (A; I)$ and a term $t(x_1, \dots, x_k)$, the interpretation of $t$ in $\mathcal{A}$ is a $k$-ary operation $t^I \colon A^k \to A$ defined on any $k$-tuple $(a_1, \dots, a_k) \in A^k$ as follows:

1° if $t$ is the variable $x_i$ where $i \le k$, then $t^I(a_1, \dots, a_k) = a_i$;

2° if $t$ is a constant symbol $c \in C$, then $t^I(a_1, \dots, a_k) = c^I$;

3° if $t$ is of the form $f(t_1, \dots, t_n)$ where $f \in F$, $\nu(f) = n$, and the interpretations $t_j^I$ of the terms $t_1(x_1, \dots, x_k)$, $\dots$ , $t_n(x_1, \dots, x_k)$ are already defined, then

$$t^I(a_1, \dots, a_k) = f^I\big(t_1^I(a_1, \dots, a_k), \dots, t_n^I(a_1, \dots, a_k)\big)$$

In particular, the interpretation of a constant term $t$ is always an element $t^I \in A$. Instead of $t^I$ we frequently write $t^{\mathcal{A}}$ or just $t$ for any term $t$.

*Formulas* of a first order language $L$, briefly $L$-formulas are built of atomic formulas by means of logical connectives and quantifiers. *Atomic formulas* of the language $L$ are expressions of the form $t_1 = t_2$ where $t_1$, $t_2$ are arbitrary $L$-terms, and $r(t_1, \dots, t_n)$ where $r \in R$ is a relational symbol, $\nu(r) = n$, and $t_1, \dots, t_n$ are arbitrary $L$-terms (instead of $r(t_1, t_2)$ we frequently write $t_1 \, r \, t_2$). The set $\mathrm{Form}(L)$ of all $L$-terms is the smallest set such that

1° if $\varphi$ is an atomic formula then $\varphi \in \mathrm{Form}(L)$ (every atomic formula is a formula);

2° if $\varphi, \psi \in \mathrm{Form}(L)$ then $\neg\varphi, (\varphi \wedge \psi), (\varphi \vee \psi), (\varphi \Rightarrow \psi), (\varphi \Leftrightarrow \psi) \in \mathrm{Form}(L)$
(if $\varphi$, $\psi$ are $L$-formulas then so are $\neg\varphi$, $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \Rightarrow \psi)$, $(\varphi \Leftrightarrow \psi)$);

3° if $\varphi \in \mathrm{Form}(L)$ and $x$ is a variable, then $(\forall\, x)\varphi, (\exists\, x)\varphi \in \mathrm{Form}(L)$
(for any formula $\varphi$ and variable $x$, the expressions $(\forall\, x)\varphi$, $(\exists\, x)\varphi$ are formulas).

This is a *recursive definition*, again. In order to prove that some set of $L$-formulas contains all the $L$-formulas, it suffices to show that it fulfills all the three conditions 1°, 2° and 3°.

Similarly as in Propositional Calculus, we tend to omit unnecessary parentheses. On the other hand, in order to promote readability, we sometimes use parentheses not required by the above definition. For instance, the modified expressions $\neg(\varphi), (\varphi) \wedge (\psi)$, etc., could sometimes be better legible than the "rigorous" formulas $\neg\varphi$, $\varphi \wedge \psi$, etc., respectively. Atomic formulas of the form $t_1 = t_2$ are called *identities*. The conjunction of two identities $(t_1 = t_2) \wedge (t_2 = t_3)$ is frequently abbreviated to $t_1 = t_2 = t_3$. Instead of $\neg(t_1 = t_2)$ we usually write $t_1 \ne t_2$; $\neg(t_1 \, r \, t_2)$ is sometimes abbreviated to $t_1 \not r \, t_2$.

Consecutive quantifications with the same quantifier $(\mathrm{Q}\, x_1) \dots (\mathrm{Q}\, x_n)$ will be abbreviated to $(\mathrm{Q}\, x_1, \dots, x_n)$. For instance, we will write $(\forall\, x_1, \dots, x_n)(\exists\, u, v)(\forall\, z)\varphi$ instead of $(\forall\, x_1) \dots (\forall\, x_n)(\exists\, u)(\exists\, v)(\forall\, z)\varphi$.

An *occurrence of a variable* $x$ in a formula $\varphi$ is simply any occurrence of the symbol $x$ in some of the atomic formulas out of which $\varphi$ is built. Such an occurrence is called *bounded* if it falls under the range of some quantifier, otherwise it is called *free*.

**Example.** In the formula $\varphi$

$$(\forall\, x)(x + y = y + x) \wedge (\exists\, y)(\forall\, u)(x \leq y + z)$$

of the first order language containing a binary operation symbol $+$ and a binary predicate symbol $\leq$ both the occurrences of $x$ in the atomic formula $x + y = y + x$ are bounded while its occurrence in $x \leq y+z$ is free, both the occurrences of $y$ in the atomic formula $x + y = y + x$ are free, while its occurrence in $x \leq y + z$ is bounded, the occurrence of $z$ in the atomic formula $x \leq y + z$ is free, finally, the variable $u$ has no occurrence in $\varphi$.

*Sentences* or *closed formulas* are $L$-formulas containing no free variables. If all the variables occurring *freely* in a formula $\varphi$ are contained in the list $x_1, \ldots, x_n$, we write $\varphi(x_1, \ldots, x_n)$. Exception: writing $\varphi$ doesn't necessarily mean that $\varphi$ is a closed formula.

The satisfaction of an $L$-formula $\varphi(x_1, \ldots, x_n)$ by the elements $a_1, \ldots, a_n \in A$ of some $L$-structure $\mathcal{A} = (A; I)$ is denoted by $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n)$ and read as $\varphi(a_1, \ldots, a_n)$ *is satisfied* or *true in* $\mathcal{A}$. It is defined recursively:

- if $\varphi$ is $t_1 = t_2$, then $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n)$ if and only if $t_1^I(a_1, \ldots, a_n) = t_2^I(a_1, \ldots, a_n)$
- if $\varphi$ is $r(t_1, \ldots, t_m)$, then $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n)$ if and only if $\big(t_1^I(a_1, \ldots, a_n), \ldots, t_m^I(a_1, \ldots, a_n)\big) \in r^I$
- if $\varphi$ is $\neg\psi$, then $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n)$ if and only if it is not true that $\mathcal{A} \vDash \psi(a_1, \ldots, a_n)$
- if $\varphi$ is $\psi \wedge \chi$, then $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n)$ if and only if both $\mathcal{A} \vDash \psi(a_1, \ldots, a_n)$ and $\mathcal{A} \vDash \chi(a_1, \ldots, a_n)$
- if $\varphi$ is $\psi \vee \chi$, then $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n)$ if and only if $\mathcal{A} \vDash \psi(a_1, \ldots, a_n)$ or $\mathcal{A} \vDash \chi(a_1, \ldots, a_n)$ (in the nonexclusive meaning)
- if $\varphi$ is $\psi \Rightarrow \chi$, then $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n)$ if and only if from $\mathcal{A} \vDash \psi(a_1, \ldots, a_n)$ it follows that $\mathcal{A} \vDash \chi(a_1, \ldots, a_n)$
- if $\varphi$ is $\psi \Leftrightarrow \chi$, then $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n)$ if and only if the conditions $\mathcal{A} \vDash \psi(a_1, \ldots, a_n)$ and $\mathcal{A} \vDash \chi(a_1, \ldots, a_n)$ are equivalent
- if $\varphi$ is $(\forall\, x)\psi$, then $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n)$ if and only if $\mathcal{A} \vDash \psi(a, a_1, \ldots, a_n)$ for every $a \in A$
- if $\varphi$ is $(\exists\, x)\psi$, then $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n)$ if and only if $\mathcal{A} \vDash \psi(a, a_1, \ldots, a_n)$ for some $a \in A$

The above list can be given an easier to survey and remember form rewriting it using the same symbols for logical connectives, quantifiers and the relation of equality, both in the first order language $L$ we are dealing with, as well as in the common English (which is a part of our metalanguage). In order to avoid the impending confusion, such an attitude is usually introduced with the phrase “by abuse of notation”. The reader should carefully inspect the rewritten version and identify the corresponding role of every particular occurrence of the logical symbols in the new list below:

$$\begin{aligned}
\mathcal{A} \vDash (t_1 = t_2)(a_1, \ldots, a_n) &\Leftrightarrow \mathcal{A} \vDash t_1(a_1, \ldots, a_n) = t_2(a_1, \ldots, a_n) \\
\mathcal{A} \vDash r(t_1, \ldots, t_m)(a_1, \ldots, a_n) &\Leftrightarrow \mathcal{A} \vDash r(t_1(a_1, \ldots, a_n), \ldots, t_m(a_1, \ldots, a_n)) \\
\mathcal{A} \vDash \neg\varphi(a_1, \ldots, a_n) &\Leftrightarrow \neg(\mathcal{A} \vDash \varphi(a_1, \ldots, a_n)) \Leftrightarrow \mathcal{A} \nvDash \varphi(a_1, \ldots, a_n) \\
\mathcal{A} \vDash (\varphi \wedge \psi)(a_1, \ldots, a_n) &\Leftrightarrow \big(\mathcal{A} \vDash \varphi(a_1, \ldots, a_n) \wedge \mathcal{A} \vDash \psi(a_1, \ldots, a_n)\big) \\
\mathcal{A} \vDash (\varphi \vee \psi)(a_1, \ldots, a_n) &\Leftrightarrow \big(\mathcal{A} \vDash \varphi(a_1, \ldots, a_n) \vee \mathcal{A} \vDash \psi(a_1, \ldots, a_n)\big) \\
\mathcal{A} \vDash (\varphi \Rightarrow \psi)(a_1, \ldots, a_n) &\Leftrightarrow \big(\mathcal{A} \vDash \varphi(a_1, \ldots, a_n) \Rightarrow \mathcal{A} \vDash \psi(a_1, \ldots, a_n)\big) \\
\mathcal{A} \vDash (\varphi \Leftrightarrow \psi)(a_1, \ldots, a_n) &\Leftrightarrow \big(\mathcal{A} \vDash \varphi(a_1, \ldots, a_n) \Leftrightarrow \mathcal{A} \vDash \psi(a_1, \ldots, a_n)\big) \\
\mathcal{A} \vDash (\forall\, x)\varphi(x, a_1, \ldots, a_n) &\Leftrightarrow (\forall\, a \in A)\big(\mathcal{A} \vDash \varphi(a, a_1, \ldots, a_n)\big) \\
\mathcal{A} \vDash (\exists\, x)\varphi(x, a_1, \ldots, a_n) &\Leftrightarrow (\exists\, a \in A)\big(\mathcal{A} \vDash \varphi(a, a_1, \ldots, a_n)\big)
\end{aligned}$$

For $\varphi(x_1, \ldots, x_n, y_1, \ldots, y_m)$ and $a_1, \ldots, a_n \in A$ we write $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n, y_1, \ldots, y_m)$ if and only if $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n, b_1, \ldots, b_m)$ for all $b_1, \ldots, b_m \in A$, i.e., if and only if $\mathcal{A} \vDash (\forall\, y_1, \ldots, y_m)\varphi(a_1, \ldots, a_n, y_1, \ldots, y_m)$. In particular, $\mathcal{A} \vDash \varphi(x_1, \ldots, x_n)$ means that $\mathcal{A} \vDash \varphi(a_1, \ldots, a_n)$ for all $a_1, \ldots, a_n \in A$, i.e., $\mathcal{A} \vDash (\forall\, x_1, \ldots, x_n)\varphi(x_1, \ldots, x_n)$.

## First Order Theories and Models

A *first order theory* $T$, i.e., a theory in a first order language $L$, is represented by and identified with the set of its *specific axioms* $T \subseteq \mathrm{Form}(L)$. An $L$-structure $\mathcal{A}$ is said to *satisfy* the theory $T$ or to be a *model* of $T$ if $\mathcal{A}$ satisfies all the axioms of $T$. We write $\mathcal{A} \vDash T$; thus we have

$$\mathcal{A} \vDash T \quad \text{if and only if} \quad \mathcal{A} \vDash \varphi \text{ for every } \varphi \in T$$

We say that a formula $\psi$ is a *logical consequence* (of the axioms) of the first order theory $T$, or that $\psi$ is *true* in $T$ if $\psi$ is true in every model $\mathcal{A}$ of the theory $T$. In that case we write $T \vDash \psi$. Thus we have

$$T \vDash \psi \quad \text{if and only} \quad \mathcal{A} \vDash \psi \text{ for every model } \mathcal{A} \vDash T$$

Our goal is to describe the semantic notion of truth or satisfaction or logical consequence in terms of the syntactic notion of *provability.* This will take part in the next section. However, before turning our attention to that point, it is desirable to get some acquaintance with a handful of important examples of first order theories and their models.

Preliminarily, we can divide first order theories into the following two categories:

(a) First order theories describing a variety of different structures sharing the same first order language and singled out by some common properties formulated as axioms of the corresponding theory. Some subclasses of the class of all models of that theory can be described in terms of some additional axioms, as well as by some properties not formulated in terms of that first order language. From the theories listed bellow the Theory of Groups, various Theories of Rings and Fields, Vector Spaces, Theories of Order, Boolean Algebras and the Theories of Ordered Rings and Fields belong to this family.

(b) First order theories attempting to describe a single mathematical structure "as completely as possible". As we shall see later on, such attempts are unattainable, except for some trivial cases. Our first example of this kind is furnished by Peano Arithmetic, aiming to fully describe the structure of all natural numbers with the addition and multiplication. The second example includes the Zermelo-Fraenkel Set Theory with the Axiom of Choice (in a not quite precisely delineated version) which should grasp the Universe of Sets.

**Groups.** The *Theory of Groups* or *Group Theory* has several alternative axiomatizations in slightly differing first order languages. A *group* is simply a model of Group Theory.

(a) In the first order language containing a binary operation symbol $\cdot$ (multiplication), a constant symbol $e$ (unit or neutral element) and a unary operation symbol $^{-1}$ (taking inverses), the axioms of the Theory of Groups are formed by the following identities:

$$x \cdot (y \cdot z) = (x \cdot y) \cdot z$$
$$x \cdot e = x = e \cdot x$$
$$x \cdot x^{-1} = e = x^{-1} \cdot x$$

expressing the associativity of the multiplication and the facts that $e$ is its unit element and that $x^{-1}$ is the inverse element of $x$. Then groups are structures $\big(G;\, \cdot, e, {}^{-1}\big)$ satisfying the just stated axioms.

A group is called *commutative* or *abelian* if it satisfies the commutative law $x{\cdot}y = y{\cdot}x$. Group Theory is sometimes, especially in the abelian case, formulated in the language using the binary operation symbol $+$ (addition), the constant symbol $0$ (zero) and the unary operation symbol $-$ (minus, the inverse element $-x$ is called the opposite element to $x$).

(b) Omitting the unary operation symbol $^{-1}$ from the language of Group Theory, the last axiom, expressing the existence of inverse elements, has to be formulated in a slightly more complicated way

$$(\forall\, x)(\exists\, y)(x \cdot y = e = y \cdot x)$$

Then groups are considered as structures $(G;\, \cdot, e)$ satisfying the corresponding three axioms.

(c) In the language containing just the symbol of multiplication, the axioms expressing the existence of the unit element and of the inverse elements are merged into a single more complex axiom

$$(\exists\, u)(\forall\, x)\big(x \cdot u = x = u \cdot x \wedge (\exists\, y)(x \cdot y = u = y \cdot x)\big)$$

Another possibility is represented by the axiom

$$(\forall\, x)(\exists\, y)(\forall\, z)\big(z \cdot (x \cdot y) = z = (y \cdot x) \cdot z\big)$$

Then a group is a structure $(G;\, \cdot)$ satisfying the associative law and one (hence both) of the last two axioms.

It is clear that a group $\left(G;\, \cdot, e, {}^{-1}\right)$ in the sense of (a) can be made a group in the sense of (b) or (c) by omitting the interpretations of the superfluous symbols. The other way round, for Group Theory in the sense of (c) one can extend its language by the missing symbols and define the unit element and the inverse element operation by

$$\begin{aligned} u = e \;&\Leftrightarrow\; (\forall\, x)(x \cdot u = x = u \cdot x) \\ y = x^{-1} \;&\Leftrightarrow\; x \cdot y = e = y \cdot x \end{aligned}$$

respectively. Then a group $(G;\, \cdot)$ in the sense of (c) becomes a group in the sense of (b) or (a).

Some examples of commutative groups (in the language with a single binary operation) are the additive groups $(\mathbb{Z};\, +)$ of integers, $(\mathbb{Z}_n;\, +)$ of remainders modulo $n \geq 2$, of rationals $(\mathbb{Q};\, +)$, etc. Some examples of noncommutative groups are provided by the structures $(S(X);\, \circ)$ of all bijective maps (permutations) of any set $X$ with more than two elements into itself and the operation of composition, or by the structures $(\mathrm{GL}(n, \mathbb{R});\, \cdot)$ of all invertible real $n \times n$ matrices, for $n \geq 2$, with the operation of matrix multiplication.

**Rings and Fields.** (a) The *Theory of Rings* or *Ring Theory* is formulated in the language with two binary operation symbols $+$ (addition), $\cdot$ (multiplication) and a constant symbol 0 (zero). Then a ring is a structure $(A;\, +, \cdot, 0)$ satisfying the axioms of Ring Theory. The axioms express that $(A;\, +, 0)$ is an abelian group, the associative law for multiplication and two distributive laws

$$x \cdot (y + z) = (x \cdot y) + (x \cdot z) \qquad (x + y) \cdot z = (x \cdot z) + (y \cdot z)$$

usually written simply as $x(y + z) = xy + xz$ and $(x + y)z = xz + yz$. Models of Ring Theory are called *rings*.

A ring is called *commutative* if it satisfies the commutative law for multiplication $x \cdot y = y \cdot x$. A commutative ring is called an *integral domain* if satisfies the axiom

$$xy = 0 \;\Rightarrow\; (x = 0 \vee y = 0)$$

(b) The*Theory of Rings with Unit* or the *Theory of Unitary Rings* is formulated in the language obtained by extending the language of Ring Theory by a new constant symbol 1, denoting the unit of multiplication and adding the identities $x \cdot 1 = x = 1 \cdot x$ to the axioms of Ring Theory.

(c) The *Theory of Fields* is obtained from the Theory of Commutative Rings with Unit by adding to it the axiom $0 \neq 1$ and the axiom

$$(\forall\, x)(x \neq 0 \;\Rightarrow\; (\exists\, y)(x \cdot y = 1))$$

stating the existence of multiplicative inverses for all nonzero elements. A *field* is simply a model of the Theory of Fields.

(d) The *Theory of Division Rings* is obtained from the Theory of Unitary Rings by extending it by the condition $0 \neq 1$ and the axiom of inverses

$$(\forall\, x)(x \neq 0 \;\Rightarrow\; (\exists\, y)(x \cdot y = 1 = y \cdot x))$$

A *division ring* is simply a model of this theory.

Even integers form a commutative ring $(2\mathbb{Z};\, +, \cdot, 0)$ without unit. The integers and the remainders modulo $n \geq 2$ form commutative rings with unit $(\mathbb{Z};\, +, \cdot, 0, 1)$ and $(\mathbb{Z}_n;\, +, \cdot, 0, 1)$, respectively. Examples of noncommutative rings with unit are provided by the structures $\big(\mathbb{R}^{n\times n};\, +, \cdot, 0, I\big)$ of all real $n \times n$ matrices, for $n \geq 2$, with operations of addition and multiplication of matrices, the zero matrix $0$ and the unit matrix $I$. Examples of fields are the structures $(\mathbb{Q};\, +, \cdot, 0, 1)$, $(\mathbb{R}; +, \cdot, 0, 1)$, $(\mathbb{C};\, +, \cdot, 0, 1)$ of rational, real and complex numbers, respectively, as well as the structures $(\mathbb{Z}_p;\, +, \cdot, 0, 1)$ of remainders modulo any prime number $p$. Clearly, every field is an integral domain; however, the integers $(\mathbb{Z};\, +, \cdot, 0, 1)$ form an integral domain which is not a field. An example of a non commutative division ring, i.e., a division ring which is not a field, is provided by the *quaternions* $(\mathbb{H};\, +, \cdot, 0, 1)$. Quaternions represent a four dimensional version of complex numbers, i.e., they are numbers of the form $q_0 + q_1\,\mathrm{i} + q_1\,\mathrm{j} + q_3\,\mathrm{k}$, where i, j, k are three imaginary units, satisfying $\mathrm{i}^2 = \mathrm{j}^2 = \mathrm{k}^2 = \mathrm{i}\,\mathrm{j}\,\mathrm{k} = -1$. Their equality and addition are defined componentwise, while their multiplication is defined in the only possible way extending the above relations between the generators i, j, k enforced by the axioms of unitary rings.

(e) A field $(F;\, +, \cdot, 0, 1)$ is called *algebraically closed* if it satisfies the infinite list of axioms

$$(\forall\, u_1, \ldots, u_n)(\exists\, x)\big(x^n + u_1 x^{n-1} + \cdots + u_{n-1} x + u_n = 0\big)$$

postulating the existence of roots of all polynomials of any degree $n \geq 2$ with coefficients from $F$. The field $(\mathbb{C};\, +, \cdot, 0, 1)$ of all complex numbers and the field $(\mathbb{A};\, +, \cdot, 0, 1)$ of all algebraic numbers (i.e., the roots of polynomials with rational coefficients) are examples of algebraically closed fields.

(f) A field $(F;\, +, \cdot, 0, 1)$ is called *formally real* if it satisfies the infinite list of axioms

$$x_1^2 + \cdots + x_n^2 = 0 \;\Rightarrow\; x_1 = \ldots = x_n = 0$$

for every positive $n \in \mathbb{N}$, requiring that a sum of squares of nonzero elements is never 0. The *Theory of Real Closed Fields* is the extension of the *Theory of Formally Real Fields* by the axiom

$$(\forall\, x)(\exists\, y)\big(y^2 = x \,\vee\, y^2 = -x\big)$$

postulating the existence of the square root of either $x$ or $-x$ for any $x$, as well as the infinite list of axioms

$$(\forall\, u_1, \ldots, u_n)(\exists\, x)\big(x^n + u_1 x^{n-1} + \cdots + u_{n-1} x + u_n = 0\big)$$

guaranteeing the existence of at least one root for every polynomial of any *odd* degree $n \geq 3$ with coefficients from $F$. The field $(\mathbb{R}; +, \cdot, 0, 1)$ of all real numbers and the field $(\mathbb{R} \cap \mathbb{A}; +, \cdot, 0, 1)$ of all real algebraic numbers are examples of real closed fields. The field $(\mathbb{Q}; +, \cdot, 0, 1)$ of all rational numbers is formally real but not real closed.

The following, though rather familiar example is worthwhile to notice since it shows that the sets of functional or relational symbols of a first order language may themselves carry their own first order structure.

**Vector spaces over a field.** For a fixed field $(F; +, \cdot, 0, 1)$ we introduce the first order language $L(F)$ which has no relational symbols, a single constant symbol $\mathbf{0}$, a binary operation symbol $+$ and the set $F$ of unary operation symbols. A typical structure of the language $L(F)$ is denoted as $\mathcal{V} = (V; F, +, \mathbf{0})$; the elements of the set $V$ are called *vectors*. The elements $f \in F$ in role of unary operation symbols are referred to as *scalars*; instead of $f(x)$ we usually write just $fx$ and this result is referred as the *scalar multiple* of $x$ by $f$. *Vector spaces* over the field $F$ are structures $\mathcal{V} = (V; F, +, \mathbf{0})$ of the language $L(F)$ satisfying the axioms expressing that $(V; +, \mathbf{0})$ is an abelian group, as well as the axioms

$$\begin{aligned} 1x &= x \\ f(x+y) &= fx + fy \\ (fg)x &= f(gx) \\ (f+g)x &= fx + gx \end{aligned}$$

for any scalars $f, g \in F$. The reader should realize that the last three equalities are in fact *axiom schemes* and not single axioms.

**Theories of Order.** The *Theory of Partial Order* is formulated alternatively in the first order language with a single binary relational symbol $<$ (strict partial order) or $\leq$ (non-strict partial order). The strict version is given by the axioms

$$\neg(x < x) \qquad \neg(x < y \wedge y < x) \qquad (x < y \wedge y < z) \Rightarrow x < z$$

The non-strict (and more frequently used version) has the axioms

$$x \leq x \qquad (x \leq y \wedge y \leq x) \Rightarrow x = y \qquad (x \leq y \wedge y \leq z) \Rightarrow x \leq z$$

A *partially ordered set* (*poset*) is simply a model $(P; <)$ or $(P; \leq)$ of the corresponding version of the theory.

Obviously, a poset $(P; <)$ can be converted into a poset $(P; \leq)$ defining the non-strict partial order by

$$x \leq y \Leftrightarrow x < y \vee x = y$$

Vice versa, a poset $(P; \leq)$ can be made to a poset $(P; <)$ defining the strict partial order by

$$x < y \Leftrightarrow x \leq y \wedge x \neq y$$

The reader should be able to switch between the two versions anytime.

The *Theory of Ordered Sets*, sometimes called also the *Theory of Total Order* or the *Theory of Linear Order*, is obtained by adding the *trichotomy axiom*

$$x < y \lor x = y \lor y < x$$

or the *dichotomy axiom*

$$x \le y \lor y \le x$$

respectively, to the corresponding version of the Theory of Partial Order.

**Boolean Algebras.** The language of the *Theory Boolean Algebras* has two binary operation symbols $\land$ (meet), $\lor$ (join), one unary operation symbol $'$ (complement) and two constant symbols 0 and 1. This violates the implicit convention that the specific symbols of any first order language should be clearly distinguished from its logical symbols. However, since all the axioms of the Theory of Boolean Algebras are identities and do not contain any logical connectives, there is no danger of confusion. On the other hand, this notation points to the familiar connection between Boolean algebras and Propositional Calculus. A *Boolean algebra* $\mathcal{B} = (B; \land, \lor, ', 0, 1)$ is simply a model of the theory with the following axioms:

| | | |
|---|---|---|
| $x \land y = y \land x$ | $x \lor y = y \lor x$ | (commutative laws) |
| $x \land (y \land z) = (x \land y) \land z$ | $x \lor (y \lor z) = (x \lor y) \lor z$ | (associative laws) |
| $x \land x = x$ | $x \lor x = x$ | (idempotent laws) |
| $x \land (x \lor y) = x$ | $x \lor (x \land y) = x$ | (absorbtion laws) |
| $x \land (y \lor z) = (x \land y) \lor (x \land z)$ | $x \lor (y \land z) = (x \lor y) \land (x \lor z)$ | (distributive laws) |
| $(x \land y)' = x' \lor y'$ | $(x \lor y)' = x' \land y'$ | (de Morgan laws) |
| $x \land 0 = 0$ | $x \lor 0 = x$ | (laws of 0) |
| $x \land 1 = x$ | $x \lor 1 = 1$ | (laws of 1) |
| $x \land x' = 0$ | $x \lor x' = 1$ | (laws of complement) |

This axiom list is redundant: some of its items can be derived as consequences of the remaining ones. A partial order on every Boolean algebra can be introduced via any of the following two equivalent conditions:

$$x \le y \Leftrightarrow x = x \land y \Leftrightarrow x \lor y = y$$

Then 0 and 1 become the smallest and the biggest element, respectively, with respect to this order in $\mathcal{B}$.

The typical examples of Boolean algebras are formed by the power sets of any sets. More precisely, for every set $I$ its power set $\mathcal{P}(I) = \{X : X \subseteq I\}$ with the operations of set-theoretical intersection, union and complement with respect to the set $I$, the empty set $\emptyset$ as 0 and the whole set $I$ as 1, gives rise to the Boolean algebra $(\mathcal{P}(I); \cap, \cup, ', \emptyset, I)$. The order relation $X \le Y$ in $\mathcal{P}(I)$ coincides with the set-theoretical inclusion $X \subseteq Y$.

Another example of a Boolean algebra can be obtained from the set $\mathrm{VF}(P)$ of all propositional forms over any set of propositional variables $P \neq \emptyset$ interpreting the equality relation as the relation of logical equivalence $A \equiv B$, with logical connectives $\wedge$, $\vee$ and $\neg$ in the role of the Boolean operations of meet, join and complement, respectively, and with any tautology in the role of $1$ and any contradiction in the role of $0$.

**Exercise.** (a) Show that both the conditions defining the relation of partial order $\leq$ on Boolean algebras are indeed equivalent, and that $\leq$ defined that way is indeed a (non strict) partial order.

(b) Show that the meet $x \wedge y$ is the infimum and the join $x \vee y$ is the supremum of the set $\{x, y\}$ with respect to the partial order $\leq$, respectively. This is to say that for any $z$ we have $z \leq x$ and $z \leq y$ if and only if $z \leq x \wedge y$, and, similarly, $x \leq z$ and $y \leq z$ if and only if $x \vee y \leq z$.

(c) Prove the *law of double complement* $x'' = x$, as well as the identities $0' = 1$, $1' = 0$ from the axioms of Boolean algebras.

**Ordered Rings and Fields.** The *Theory of Ordered Rings* is obtained by extending the language of Ring Theory by the binary relational symbol $<$ and adding to its axioms the strict version of the axioms of (total) order, as well as the axioms

$$x < y \Rightarrow x + z < y + z$$

$$(x < y \wedge 0 < z) \Rightarrow (xz < yz \wedge zx < zy)$$

expressing that the operations of the addition of any element $z$ as well as the multiplication by any positive element $z$ are increasing.

The same procedure applies to the Theory of Unitary Rings, the Theory of Commutative Rings, the Theory of Fields, etc., yielding the *Theory Ordered Unitary Rings*, the *Theory of Ordered Commutative Rings*, the *Theory of Ordered Fields*, etc., respectively. The integers $(\mathbb{Z};\, +, \cdot, 0, 1, <)$ provide a representative (and minimal) example of an ordered commutative ring with unit. The rationals $(\mathbb{Q};\, +, \cdot, 0, 1, <)$ and the reals $(\mathbb{R};\, +, \cdot, 0, 1, <)$ form ordered fields. It can be proved that neither the field of complex numbers $(\mathbb{C};\, +, \cdot, 0, 1)$ nor any finite field, e.g., the fields $(\mathbb{Z}_p;\, +, \cdot, 0, 1)$ of remainders modulo a prime $p$, can be turned into an ordered field by any ordering relation $<$.

An ordered field $(F;\, +, \cdot, 0, 1, <)$ is called *real closed* if it is a formally real field satisfying the condition

$$x \geq 0 \;\Rightarrow\; (\exists\, z)\big(x = z^2\big)$$

and the infinite list of conditions stating that every polynomial of *odd* degree $n \geq 3$ with coefficients from $F$ has at least one root in $F$, similarly as in the case of (unordered) real closed fields. The *Theory of Ordered Real Closed Fields* is obtained as the extension of the *Theory of Ordered Fields* by these axioms. It can be easily verified that every real closed field $(F;\, +, \cdot, 0, 1)$ can be turned into an ordered field $(F;\, +, \cdot, 0, 1, <)$ by defining the (nonstrict) order relation on it as follows:

$$x \leq y \;\Leftrightarrow\; (\exists\, z)\big(y = x + z^2\big)$$

The other way round, for every ordered real closed field $(F;\, +, \cdot, 0, 1, <)$, its reduct $(F;\, +, \cdot, 0, 1)$, obtained by omitting the order relation $<$ is a real closed field.

**Peano Arithmetic.** *Peano Arithmetic* PA is the most common first order theory describing the structure of natural numbers. In logical texts it is usually formulated in the first order language with a unary operation symbol $S$ (successor operation, i.e., adding 1), two binary operations of addition $+$ and multiplication $\cdot$, and a constant symbol 0. However, we find it more convenient to replace the successor symbol $S$ by the constant symbol 1; that way our formulation of PA will use the familiar language of the Theory of Unitary Rings (then the successor operation can be defined by $S(x) = x + 1$).

The axioms of PA can be divided into four groups. The first group consists of three axioms for the successor operation (the left column), the second group is in fact the recursive definition of addition in terms of successor (the middle column), and the third group is the recursive definition of multiplication in terms of addition (the right column):

$$\begin{array}{lll} 0 + 1 = 1 & x + 0 = x & x \cdot 0 = 0 \\ 0 \neq x + 1 & x + (y + 1) = (x + y) + 1 & x \cdot (y + 1) = (x \cdot y) + x \\ x + 1 = y + 1 \Rightarrow x = y & & \end{array}$$

The fourth group consists of infinitely many axioms comprised in the *Scheme of Induction*

$$\big(\varphi(0, \vec{u}) \wedge (\forall\, x)(\varphi(x, \vec{u}) \Rightarrow \varphi(x + 1, \vec{u}))\big) \Rightarrow (\forall\, x)\varphi(x, \vec{u})$$

where $\varphi(x, u_1, \ldots, u_n)$ is any formula in the language of PA, abbreviated to $\varphi(x, \vec{u})$.

The language of PA is usually extended by the ordering symbol $\leq$ defined by

$$x \leq y \Leftrightarrow (\exists\, z)(y = x + z)$$

and the axioms of PA imply that this is a (non-strict) linear order. Then the Scheme of Induction can be equivalently expressed in form of the *Well Ordering Principle*

$$(\exists\, x)\psi(x, \vec{u}) \Rightarrow (\exists\, x)\big(\psi(x, \vec{u}) \wedge (\forall\, y)(\psi(y, \vec{u}) \Rightarrow x \leq y)\big)$$

for any formula $\psi(x, \vec{u})$ in the language of PA. Informally, this principle expresses the condition that every nonempty set of the form $\{x\colon \psi(x, \vec{u})\}$ has the least element.

The "usual" natural numbers form the so called *standard model* $(\mathbb{N};\, +, \cdot, 0, 1)$ of PA. Every element $n \in \mathbb{N}$ is the interpretation of some constant term in the language of PA. The canonical representatives of particular natural numbers are defined recursively as follows:

1° The natural number 0 coincides with the constant symbol 0.

2° If the natural number $n$ coincides with the constant term $t$, then the natural number $n + 1$ coincides with the constant term $t + 1$.

By abuse of notation we can write $0 = 0$, $1 = 0 + 1$, $2 = (0 + 1) + 1$, $3 = ((0 + 1) + 1) + 1$, $\ldots$, $n = (\ldots((0 + 1) + 1) \cdots + 1) + 1$ (with $n$ instances of 1), etc. Thus the natural number $n$ is represented by the constant term obtained as the $n^{\text{th}}$ iterate of the successor operation applied to the constant symbol 0.

Later on we shall see that PA has some nonstandard models, as well.

**Set Theory.** Most versions of the Set Theory are formulated in the first order language with a single binary relational symbol $\in$ denoting the membership relation. The formula $x \in X$ means that $x$ *is an element of* $X$ or that $x$ *belongs to the set* $X$. The common core of these versions consists of the following four axioms:

$$\begin{aligned}
&X = Y \;\Leftrightarrow\; (\forall\, z)(z \in X \;\Leftrightarrow\; z \in Y)\\
&(\forall\, x, y)(\exists\, Z)(\forall\, z)(z \in Z \;\Leftrightarrow\; (z = x \;\vee\; z = y))\\
&(\forall\, X)(\exists\, U)(\forall\, u)(u \in U \;\Leftrightarrow\; (\exists\, x \in X)(u \in x))\\
&(\forall\, X)(\exists\, Y)(\forall\, y)(y \in Y \;\Leftrightarrow\; (\forall\, x)(x \in y \;\Rightarrow\; x \in X))
\end{aligned}$$

called the *Axiom of Extensionality*, the *Axiom of Pair*, the *Axiom of Union* and the *Power Set Axiom*, respectively, and of the following infinite list of axioms

$$(\forall\, X)(\exists\, Y)(\forall\, x)\big(x \in Y \;\Leftrightarrow\; (x \in X \;\wedge\; \varphi(x, \vec{u}))\big)$$

for any set-theoretical formula $\varphi(x, u_1, \ldots, u_n)$, called the *Scheme of Comprehension.*

The Axiom of Extensionality states that two sets $X$ and $Y$ are equal if and only if they contain the same elements. The Axiom of Pair postulates the existence of the set

$$Z = \{x, y\} = \{z \colon z = x \;\vee\; z = y\}$$

for any pair of elements $x$, $y$. The Axiom of Union guarantees the existence of the union

$$U = \bigcup X = \{u \colon (\exists\, x \in X)(u \in x)\}$$

of all the sets $x$ from a given set $X$. The Power Set Axiom postulates the existence of the power set (i.e., the set of all subsets)

$$Y = \mathcal{P}(X) = \{y \colon y \subseteq X\} = \{y \colon (\forall\, x)(x \in y \;\Rightarrow\; x \in X)\}$$

of any set $X$. Finally, the Scheme of Comprehension guarantees the possibility to single out every subset of the form

$$Y = \{x \in X \colon \varphi(x, \vec{u})\} = \{x \colon x \in X \;\wedge\; \varphi(x, \vec{u})\}$$

from a given set $X$ by means of any set-theoretical formula $\varphi(x, u_1, \ldots, u_n)$.

The *Zermelo-Fraenkel Set Theory* ZF is obtained by adding to this list the *Axiom of Foundation*, the *Axiom of Infinity* and the *Scheme of Replacement.* The *Zermelo-Fraenkel Set Theory with Choice* ZFC, which is the most common version of Set Theory used in modern mathematics, is obtained from ZF by adding to it the *Axiom of Choice* (AC). We do not include the formulation of these higher axioms of Set Theory into our elementary text. Let us just confine to the following four informal formulations: The Axiom of Foundation states that every set of sets contains as an element a set disjoint from this set. The Axiom of Infinity postulates the existence of an infinite set. The Scheme of Replacement generalizes the Scheme of Comprehension by guaranteeing even the sethood of certain images of subsets of a given set defined by set-theoretical formulas. Finally, the Axiom of Choice guarantees, for every set $X$ of pairwise disjoint nonempty sets, the existence of a set containing exactly one element from each of the sets $x \in X$.

### Axiomatization of First Order Logic and the Soundness Theorem

Similarly to Propositional Calculus, we prefer to have a brief and concise axiomatization of First Order Logic. Therefore we will proceed as if the set $\mathrm{Form}(L)$ of all formulas of any first order language $L$ were built of the atomic formulas by means of the logical connectives $\neg$ and $\Rightarrow$, and the universal quantifier $\forall$, only. By $\varphi \approx \psi$ we express that the characters $\varphi$ and $\psi$ denote the same formula. Again, the symbol $\approx$ does not belong to our first order language, similarly as the symbols $\varphi$, $\psi$, $\chi$, $L$, etc. They are symbols of our *metalanguage* by means of which we describe Predicate Calculus.

The remaining logical connectives and the existential quantifier can be introduced as the abbreviations

$$\begin{aligned}
(\varphi \wedge \psi) &\approx \neg(\varphi \Rightarrow \neg\psi), \\
(\varphi \vee \psi) &\approx (\neg\varphi \Rightarrow \psi), \\
(\varphi \Leftrightarrow \psi) &\approx (\varphi \Rightarrow \psi) \wedge (\psi \Rightarrow \varphi), \\
(\exists\, x)\varphi &\approx \neg(\forall\, x)\neg\varphi.
\end{aligned}$$

*Logical axioms* of Predicate Calculus are divided into three groups: *propositional axioms*, *quantifier axioms* and *axioms of equality*.

**Propositional axioms.** (4 axiom schemes)
For any formulas $\varphi$, $\psi$, $\chi$ the following formulas are axioms:
(PrAx 1) $\varphi \Rightarrow (\psi \Rightarrow \varphi)$
(PrAx 2) $(\varphi \Rightarrow (\psi \Rightarrow \chi)) \Rightarrow ((\varphi \Rightarrow \psi) \Rightarrow (\varphi \Rightarrow \chi))$
(PrAx 3) $(\varphi \Rightarrow \psi) \Rightarrow ((\varphi \Rightarrow \neg\psi) \Rightarrow \neg\varphi)$
(PrAx 4) $\neg\neg\varphi \Rightarrow \varphi$

If $\varphi$ is a formula and $t$ is a term, then $\varphi(t/x)$ denotes the formula obtained by the *substitution* of the term $t$ for the variable $x$. It means that every *free* occurrence of the variable $x$ in $\varphi$ is replaced by $t$. Similarly we can introduce multiple substitutions $\varphi(t_1/x_1, \ldots, t_n/x_n)$. If there's no danger of confusion then we write just $\varphi(t)$ and $\varphi(t_1, \ldots, t_n)$.

The substitution of $t$ for $x$ in $\varphi$ is *admissible* if no variable of the term $t$ falls under the range of some quantifier in $\varphi$ after substituting $t$ for $x$ in $\varphi$. Informally this means that "$\varphi(t/x)$ is telling of $t$ the same thing as $\varphi$ is telling of $x$".

**Example.** Within the integers the formula $\varphi(x) \approx (\exists\, y)(x = y + y)$ tells that $x$ is an even number. If $t$ is the term $u+x$ then $\varphi(t/x)$ is the formula $(\exists\, y)(u+x = y+y)$, telling that $u + x$ is even; this is an admissible substitution. If $t$ is the term $y$ then $\varphi(t/x)$ is the sentence $(\exists\, y)(y = y + y)$ expressing the (true) fact that the equation $y = y + y$ has some solution; this substitution is not admissible.

**Quantifier axioms.** (2 axiom schemes)
For any formulas $\varphi$, $\psi$ and any term $t$ the following formulas are axioms:

(QAx 1) $(\forall x)(\varphi \Rightarrow \psi) \Rightarrow (\varphi \Rightarrow (\forall x)\psi)$
(whenever the variable $x$ has no free occurrence in $\varphi$)
(QAx 2) $(\forall x)\varphi \Rightarrow \varphi(t/x)$
(whenever the substitution of $t$ for $x$ in $\varphi$ is admissible)

**Axioms of equality.** (3 axioms + 2 axiom schemes)
For any $n$-ary functional symbol $f$ and any $n$-ary relational symbol $r$ the following formulas are axioms:

(EAx 1) $x = x$
(EAx 2) $x = y \Rightarrow y = x$
(EAx 3) $x = y \Rightarrow (y = z \Rightarrow x = z)$
(EAx 4) $x_1 = y_1 \Rightarrow (\ldots \Rightarrow (x_n = y_n \Rightarrow f(x_1, \ldots, x_n) = f(y_1, \ldots, y_n)) \ldots)$
(EAx 5) $x_1 = y_1 \Rightarrow (\ldots \Rightarrow (x_n = y_n \Rightarrow r(x_1, \ldots, x_n) \Rightarrow r(y_1, \ldots, y_n)) \ldots)$

**Deduction rules.** Modus Ponens (MP) and Rule of Generalization (Gen)

(MP) $\dfrac{\varphi,\ \varphi \Rightarrow \psi}{\psi}$ (from $\varphi$ and $\varphi \Rightarrow \psi$ infer $\psi$)

(Gen) $\dfrac{\varphi}{(\forall x)\varphi}$ (from $\varphi$ infer $(\forall x)\varphi$)

**Exercise 1.** (a) Show that all the logical axioms are satisfied in every $L$-structure $\mathcal{A}$.
(b) Show that the deduction rule Modus Ponens is correct in the following sense: For every $L$-structure $\mathcal{A}$ and any $L$-formulas $\varphi$, $\psi$, if $\mathcal{A} \vDash \varphi$ and $\mathcal{A} \vDash \varphi \Rightarrow \psi$ then $\mathcal{A} \vDash \psi$.
(c) Show that the Rule of Generalization is correct in the following sense: For every $L$-structure $\mathcal{A}$ and any $L$-formula $\varphi$, if $\mathcal{A} \vDash \varphi$ then $\mathcal{A} \vDash (\forall x)\varphi$ for any variable $x$, no matter whether $x$ is free in $\varphi$ or not.

A *proof* in a first order theory $T$ is a finite sequence $\varphi_0, \varphi_1, \ldots, \varphi_n$ of formulas such that every item $\varphi_k$ is either a logical axiom, or a specific axiom of the theory $T$ (i.e., $\varphi_k \in T$), or it follows from the previous items by modus ponens (MP) (i.e., there are $i, j < k$ such that $\varphi_j$ has the form $\varphi_i \Rightarrow \varphi_k$) or by the rule of generalization (Gen) (i.e., there is a $j < k$ such that $\varphi_k$ has the form $(\forall x)\varphi_j$ for some variable $x$).

A formula $\psi$ is *provable* in a theory $T$ if there is a proof $\varphi_0, \varphi_1, \ldots, \varphi_n$ in $T$ such that its last item $\varphi_n$ coincides with $\psi$. In symbols, $T \vdash \psi$. Instead of $\emptyset \vdash \psi$ we write just $\vdash \psi$; it means that $\psi$ is provable from the logical axioms, only.

**Exercise 2.** Show that the following first order schemes are provable just from the logical axioms:

(a) all propositional tautologies
(b) $\varphi(t/x) \Rightarrow (\exists x)\varphi$ (if the substitution $t/x$ in $\varphi$ is *admissible*)
(c) $(x_1 = y_1 \wedge \ldots \wedge x_n = y_n) \Rightarrow t(x_1, \ldots, x_n) = t(y_1, \ldots, y_n)$ (for any term $t$)
(d) $(x_1 = y_1 \wedge \ldots \wedge x_n = y_n) \Rightarrow (\varphi(x_1, \ldots, x_n) \Leftrightarrow \varphi(y_1, \ldots, y_n))$
(for any formula $\varphi$ such that all the substitutions $y_i/x_i$, in $\varphi$ are admissible)

**Soundness Theorem.** *Let $T$ be a theory in a first order language $L$. Then, for every $L$-formula $\psi$, if $T \vdash \psi$ then $T \vDash \psi$.*

*Demonstration.* Let $T \vdash \psi$ and $\varphi_0, \varphi_1, \ldots, \varphi_n$ be a proof of $\psi$ in $T$. We will show that $\mathcal{A} \vDash \varphi_k$, for any model $\mathcal{A} \vDash T$ of the theory $T$ and *each* $k \leq n$. Then, of course, $\mathcal{A} \vDash \psi$, since $\psi$ is $\varphi_n$. Each $\varphi_k$ is either a logical axiom, in which case $\mathcal{A} \vDash \varphi_k$ for every $L$-structure $\mathcal{A}$, or a specific axiom of $T$, in which case $\mathcal{A} \vDash \varphi_k$ as $\mathcal{A} \vDash T$, or $\varphi_k$ follows from some previous proof items by (MP) or by (Gen). In the (MP) case there are $i, j < k$ such that $\varphi_j$ has the form $\varphi_i \Rightarrow \varphi_k$. Now, for any $L$-structure $\mathcal{A}$, assuming that we already have $\mathcal{A} \vDash \varphi_i$ and $\mathcal{A} \vDash \varphi_j$, i.e., $\mathcal{A} \vDash \varphi_i \Rightarrow \varphi_k$, we can conclude $\mathcal{A} \vDash \varphi_k$. In the (Gen) case there is a $j < k$ such that $\varphi_k$ has the form $(\forall\, x)\varphi_i$. Again, for any $L$-structure $\mathcal{A}$, assuming that we already have $\mathcal{A} \vDash \varphi_j$, we can conclude $\mathcal{A} \vDash (\forall\, x)\varphi_j$, i.e., $\mathcal{A} \vDash \varphi_k$. (Cf. Exercise 1.)

Later on we will also establish the converse of the Soundness Theorem.

**Completeness Theorem.** *Let $T$ be a theory in a first order language $L$. Then, for every $L$-formula $\psi$, if $T \vDash \psi$ then $T \vdash \psi$.*

At this moment, the reader should return to the remarks following the demonstration of the Soundness Theorem and the first formulation of the Completeness Theorem in Propositional Calculus and realize that they equally apply to their first order versions.

As it follows from the following example, the fine balance between the syntax and semantics which we have both in the Propositional as well as in the First Order Logic is no way self-evident or automatic.

**Example.** (Finite Model Semantics) Let $L$ be a first order language. For any theory $T$ in $L$ and any $L$-formula $\varphi$ we define the *finite satisfaction relation* $T \vDash_{\text{fin}} \varphi$ if $T \vDash \mathcal{A}$ for every *finite* model $\mathcal{A}$ of the theory $T$, i.e., if and only if $\varphi$ is satisfied in every *finite* model of the theory $T$. By methods going beyond the scope of our elementary course it can be shown that this *Finite Model Semantics* cannot be axiomatized in a way enabling to establish both the corresponding versions of the Soundness Theorem and of the Completeness Theorem. More precisely, for any sound axiomatization, consisting of finitely many axiom schemes and finitely many rules of inference, the resulting provability relation $T \vdash_{\text{fin}} \varphi$ does not exhaust the finite satisfaction relation $T \vDash_{\text{fin}} \varphi$. This is to say that it is possible to find a theory $T$ and a sentence $\varphi$ such that $T \vDash_{\text{fin}} \varphi$, nevertheless $T \nvdash_{\text{fin}} \varphi$ for any sound provability relation $\vdash_{\text{fin}}$.

To convey to the reader at least some feeling of the issue, let us mention the deep *Wedderburn's Theorem*, stating that every finite division ring is a field. In other words, the commutative law $xy = yx$ is finitely satisfied in the theory of division rings. On the other hand, the infinite division ring of all quaternions $(\mathbb{H};\, +, \cdot, 0, 1)$ is non commutative; therefore, the commutative law for multiplication is not a first order consequence of the axioms for division rings.

**Exercise.** We say that an $L$-formula $\varphi$ is *logically valid* if it is satisfied in every $L$-structure $\mathcal{A}$. Two $L$-formulas $\varphi$, $\psi$ are called *logically equivalent*, notation $\varphi \equiv \psi$ if the formula $\varphi \Leftrightarrow \psi$ is logically valid. A formula $\varphi$ is said to be in *prenex normal form* if it has the shape $(\mathrm{Q}_1\, x_1)\dots(\mathrm{Q}_n\, x_n)\psi$ where $\mathrm{Q}_1, \dots, \mathrm{Q}_n$ are arbitrary quantifiers and $\psi$ is a quantifier-free formula.

(a) Show that, for any formulas $\varphi$, $\psi$, the following pairs of formulas are logically equivalent:

$$
\begin{aligned}
\neg(\forall\, x)\varphi &\equiv (\exists\, x)\neg\varphi & \qquad \neg(\exists\, x)\varphi &\equiv (\forall\, x)\neg\varphi \\
(\forall\, x)\varphi \wedge \psi &\equiv (\forall\, x)(\varphi \wedge \psi) & \qquad (\exists\, x)\varphi \vee \psi &\equiv (\exists\, x)(\varphi \vee \psi) \\
(\forall\, x)\varphi \Rightarrow \psi &\equiv (\exists\, x)(\varphi \Rightarrow \psi) & \qquad \psi \Rightarrow (\exists\, x)\varphi &\equiv (\exists\, x)(\psi \Rightarrow \varphi)
\end{aligned}
$$

(b) Show that, for any formulas $\varphi$, $\psi$ such that the variable $x$ has no free occurrence in $\psi$, the following pairs of formulas are logically equivalent:

$$
\begin{aligned}
(\forall\, x)\varphi \vee \psi &\equiv (\forall\, x)(\varphi \vee \psi) & \qquad (\exists\, x)\varphi \wedge \psi &\equiv (\exists\, x)(\varphi \wedge \psi) \\
(\exists\, x)\varphi \Rightarrow \psi &\equiv (\forall\, x)(\varphi \Rightarrow \psi) & \qquad \psi \Rightarrow (\forall\, x)\varphi &\equiv (\forall\, x)(\psi \Rightarrow \varphi)
\end{aligned}
$$

In each case decide which of the implications "left $\Rightarrow$ right", "right $\Rightarrow$ left" remain logically valid even without the assumption that $x$ has no free occurrence in $\psi$, and give examples that in the remaining cases this assumption cannot be omitted.

(c) Using (a) and (b) show that every $L$-formula is logically equivalent to some formula in prenex normal form.

(d) Replace any case of the logical equivalences $\theta \equiv \chi$ in (a) and (b) by the formula $\theta \Leftrightarrow \chi$ and show that all the formulas thus obtained are provable from the logical axioms (in fact just from the propositional axioms and the quantifier axioms).

## The Deduction Theorem and its Corollaries

On the way to the demonstration of the Completeness Theorem we are going to state several results which are of independent interest in their own right. The first group consists of the Deduction Theorem and its two corollaries, namely the Theorem on Proof by Contradiction and the Theorem on Proof by Distinct Cases, which are analogous to their propositional counterparts.

**Deduction Theorem.** *Let $T$ be a theory in a first order language $L$ and $\varphi$, $\psi$ be any $L$-formulas. If $\varphi$ is closed then $T \vdash \varphi \Rightarrow \psi$ if and only if $T \cup \{\varphi\} \vdash \psi$.*

The easy fact that from $T \vdash \varphi \Rightarrow \psi$ there follows $T \cup \{\varphi\} \vdash \psi$ can be established in exactly the same way as the corresponding implication in the demonstration of the Deduction Theorem in Propositional Calculus, even without the assumption that $\varphi$ is closed. It is the converse which is needed for the justification of the usual way of argumentation when proving the implication $\varphi \Rightarrow \psi$ in $T$ by proving $\psi$ in $T \cup \{\varphi\}$. This can be established in a similar, just a little bit more complicated way, as the demonstration of the corresponding implication in Propositional Calculus, again. One just has to deal additionally with the case when $\psi$ follows from some preceding item of its proof in $T \cup \{\varphi\}$ by the Rule of Generalization (Gen). Let us fill in this gap, leaving the details to the reader.

*Demonstration.* Assume that $\psi_0, \psi_1, \ldots, \psi_n$ is a proof of the formula $\psi$ in the theory $T \cup \{\varphi\}$ and $\psi_n$ follows from some previous formula $\psi_k$, where $0 \leq k < n$, by (Gen). Then $\psi_k$ is provable in $T \cup \{\varphi\}$ and, by an induction argument, we can assume that the implication $\varphi \Rightarrow \psi_k$ is provable in $T$. Thus $\psi$ has the form $(\forall\, x)\psi_k$ for some variable $x$. Now the following formulas are provable in $T$:

(1) $\varphi \Rightarrow \psi_k$
(2) $(\forall\, x)(\varphi \Rightarrow \psi_k)$ follows from (1) by (Gen)
(3) $(\forall\, x)(\varphi \Rightarrow \psi_k) \Rightarrow (\varphi \Rightarrow (\forall\, x)\psi_k)$ is an instance of the quantifier axiom scheme (QAx 1), as $\varphi$ is closed so that $x$ has no free occurrence in it
(4) $(\varphi \Rightarrow (\forall\, x)\psi_k)$ can be inferred from (2) and (3) by (MP)

Thus, finally, $T \vdash \varphi \Rightarrow \psi$.

Next we show that, in general, one cannot do without the assumption that $\varphi$ is closed.

**Example.** Let $L$ be the *language of pure equality*, i.e., the first order language without any specific symbols ($F = C = R = \emptyset$) and $T = \emptyset$ be the theory without any specific axioms in $L$. Denote by $\varphi$ the formula $x = y$ and by $\psi$ the formula $x = z$. We claim that $T \cup \{\varphi\} \vdash \psi$, nevertheless, $T \not\vdash \varphi \Rightarrow \psi$. $T \cup \{\varphi\} = \{\varphi\}$ is the theory with a single axiom $x = y$. It means that all the models of this theory have just a one-element base set. Applying (Gen) to this axiom, we see that $T \cup \{\varphi\} \vdash (\forall\, y)(x = y)$. Now, we have the quantifier axiom $(\forall\, y)(x = y) \Rightarrow x = z$, and using (MP) we infer $x = z$. At the same time, the implication $\varphi \Rightarrow \psi$, i.e., $x = y \;\Rightarrow\; x = z$ is not provable just from logical axioms, i.e., in the theory $T$. Namely, if this were the case, then it would be satisfied in every $L$-structure $\mathcal{A}$. However, every $\mathcal{A}$ with at least a two-element base set with elements $a \neq b$ violates this implication, since substituting $a$ for both $x$ and $y$ and $b$ for $z$ we have $a = a$, nevertheless $a \neq b$.

A theory $T$ in a first order language $L$ is called *inconsistent* or *contradictory* if there is some closed $L$-formula $\varphi$ such that $T \vdash \varphi$ as well as $T \vdash \neg\varphi$. Otherwise, $T$ is called *consistent* or *contradiction-free*. It can be easily verified that $T$ is inconsistent if and only if every $L$-formula is provable in $T$.

The following results follow from the Deduction Theorem in the same way as in Propositional Calculus.

**Corollary on Proof by Contradiction.** *Let $T$ be a theory in a first order language $L$ and $\varphi$ be a closed $L$-formula. Then $T \vdash \varphi$ if and only if the theory $T \cup \{\neg\varphi\}$ is contradictory (inconsistent).*

**Corollary on Proof by Distinct Cases.** *Let $T$ be a theory in a first order language $L$ and $\varphi$, $\psi$ be any $L$-formulas. If $\varphi$ is closed then $T \vdash \psi$ if and only if $T \cup \{\varphi\} \vdash \psi$ and $T \cup \{\neg\varphi\} \vdash \psi$.*

**Exercise.** Find examples showing that the assumption that $\varphi$ is closed cannot be omitted from the above Corollaries.

### Complete Theories

Another important property of first order theories closely related to consistency is that of their completeness. A theory $T$ in a first order language $L$ is called *complete* if it consistent and for any *closed* $L$-formula $\varphi$ we have $T \vdash \varphi$ or $T \vdash \neg\varphi$. In other words, $T$ is complete if and only if, for every $L$-sentence $\varphi$, either $T \vdash \varphi$ or $T \vdash \neg\varphi$ (but noth both).

**Example.** The reader may wonder why in the definition of complete theories we have required that $T$ can decide just the closed formulas, ignoring the remaining ones. For the sake of explanation, consider the formula $x = y$ and its negation $x \neq y$. If the provability of one of them were included in the requirement of completeness of a theory $T$ then, in the first case, $T$ would have just one-element models, or, in the second case, it would be contradictory. As a consequence, any consistent theory “complete” in such a sense would have trivial models, only.

Using the Theorem on Proof by Contradiction, complete theories can be characterized as maximal consistent theories in the following sense:

**Corollary on Complete Theories.** *Let $T$ be a consistent theory in a first order language $L$. Then $T$ is complete if and only if, for every $L$-sentence $\varphi$, either $T \vdash \varphi$ or the theory $T \cup \{\varphi\}$ is contradictory.*

In other words, extending the axiom list of a complete theory $T$ by any sentence $\varphi$ makes no sense: either $\varphi$ is already provable in $T$ (in which case the sets of formulas provable in $T$ and $T \cup \{\varphi\}$ coincide) or $T \cup \{\varphi\}$ turns inconsistent hence worthless.

Most of the relevant first theories occurring in mathematics are not complete. On the other hand, many of them have important complete extensions. In this place we just mention some examples of complete theories without proving their completeness.

**Divisible Abelian Groups.** The Theory of Groups is not complete: for instance, the fact that there exist both abelian as well as nonabelian groups shows that neither the commutativity law $(\forall\, x, y)(xy = yx)$ nor its negation can be proved in it. Here we describe some relatively simple complete extensions of the Theory of Abelian Groups.

An abelian group $\mathcal{G} = (G; +, 0)$ is called *nontrivial* if $(\exists\, x)(x \neq 0)$ holds in $\mathcal{G}$; it is called *divisible* if it is nontrivial and satisfies the condition

$$(\forall\, x)(\exists\, y)(n \times y = x)$$

for every integer $n \geq 2$, where $n \times x = x + \ldots + x$ with $n$-fold occurrence of $x$. $\mathcal{G}$ is called *torsion-free* if it satisfies all the conditions

$$n \times x = 0 \;\Rightarrow\; x = 0$$

for $n \geq 2$. Given a fixed $n \geq 1$, we say that $\mathcal{G}$ is a *group of exponent* $n$ if it satisfies $(\forall\, x)(n \times x = 0)$. It is known that the *Theory of Divisible Torsion-Free Abelian Groups*, as well as every *Theory of Divisible Abelian Groups of Exponent* $p$, for a fixed prime number $p$, is complete.

**Real Closed and Algebraically Closed Fields.** It can be shown that the Theory of Real Closed Fields (both in its unordered as well as in its ordered version) is complete. On the other hand, The Theory of Algebraically Closed Fields is not complete. Nonetheless, its complete extensions can be fully described.

The *characteristic* of a unitary ring $\mathcal{A} = (A;\, +, \cdot, 0, 1)$ is the least integer $\operatorname{char}(\mathcal{A}) = n \geq 1$ such that $n \times 1 = 0$, or $\operatorname{char}(\mathcal{A}) = \infty$ if $n \times 1 \neq 0$ for each $n \geq 1$ (some authors put $\operatorname{char}(\mathcal{A}) = 0$ in this case). It is known that the characteristic of any field is either a prime or $\infty$. A field $(F,\, +, \cdot, 0, 1)$ has the prime characteristic $p$ if and only if it satisfies $p \times 1 = 0$, it has the characteristic $\infty$ if and only if it satisfies all the conditions $p \times 1 \neq 0$ for every prime number $p$. Every Theory of Algebraically Closed Fields of a fixed prime Characteristic $p$, as well as the Theory of Algebraically Closed Fields of Characteristic $\infty$ is complete.

**Dense Linear Order.** A linearly ordered set $(A;\, <)$ is called *dense* if it satisfies the condition

$$(\forall\, x, y)\big(x < y \;\Rightarrow\; (\exists\, z)(x < z < y)\big)$$

$(A;\, <)$ is *without endpoints* if it satisfies the condition

$$(\forall\, x)(\exists\, y, z)(y < x < z)$$

The *Theory of Dense Linear Order without Endpoints* can be proved to be complete. Three more complete extension of the Theory of Dense Linear Order can obtained by the variation of the condition of the existence of endpoints in the obvious way.

**Atomic and Atomless Boolean Algebras.** An element $a \in B$ of a Boolean algebra $\mathcal{B} = (B;\, \wedge, \vee, {}', 0, 1)$ is called an *atom* if $a \neq 0$ and there is no element $b \in B$ such that $0 < b < a$. Formally, we can extend the language of Boolean algebras by a new unary predicate $\operatorname{At}(x)$ defined by

$$\operatorname{At}(x) \;\Leftrightarrow\; x \neq 0 \,\wedge\, (\forall\, y)(0 \leq y \leq x \;\Rightarrow\; y = 0 \,\vee\, y = x)$$

using the previously defined symbol $\leq$. Then $\mathcal{B}$ is called *atomic* if for every nonzero element of $B$ there is an atom contained in it, i.e., $\mathcal{B}$ satisfies the condition

$$(\forall\, x)(x \neq 0 \;\Rightarrow\; (\exists\, y)(\operatorname{At}(y) \,\wedge\, y \leq x))$$

$\mathcal{B}$ is called *atomless* if it has at least two elements and contains no atom. This can be expressed by the nontriviality condition $0 \neq 1$ and a kind of density axiom

$$(\forall\, x)(x \neq 0 \;\Rightarrow\; (\exists\, y)(0 < y < x))$$

The *Theory of Atomic Boolean Algebras with Infinitely Many Atoms* as well as every *Theory of Atomic Boolean Algebras with precisely $n$ Atoms* for any $n \geq 0$ are complete. Similarly, the *Theory of Atomless Boolean Algebras* is complete, too. Moreover, all the complete extensions of the Theory of Boolean Algebras can be effectively described by a pair of integer invariants; however, this description is already beyond the scope of our exposition.

**Presburger Arithmetic.** Later on, when dealing with Gödel's Incompleteness Theorems, we shall see that not only Peano Arithmetic is not complete but also its completions cannot be effectively described. On the other hand, it has an interesting complete subtheory called *Presburger Arithmetic*, describing the structure of addition (and the successor operation) of natural numbers. Its language contains just the constant symbols 0 and 1 and the operation symbol $+$; its axioms are obtained from the axioms of Peano Arithmetic by omitting those containing the symbol of multiplication, i.e., the couple forming the right most column of the seven individual axioms of PA as well as all the instances of the Scheme of Induction where the formula $\varphi(x, \vec{u})$ contains the operation symbol $\cdot$ .

## Results on Language Extensions

When proving a universally quantified statement of the form $(\forall\, x_1, \ldots, x_k)\varphi(x_1, \ldots, x_k)$ in a first order theory $T$, we usually begin with the phrase: *"Let $x_1, \ldots, x_n$ be arbitrary elements … "* That, however, means that we do not consider $x_1, \ldots, x_n$ in our proof as variables any more, and deal with them as with some unspecified constants. The following result shows that such a kind of argumentation is legitimate in First Order Logic.

**Theorem on Constants.** *Let $T$ be a theory in a first order language $L$, $\varphi(x_1, \ldots, x_k)$ be an $L$-formula and $c_1, \ldots, c_k$ be pairwise distinct constant symbols not occurring in $L$. Then*

$$T \vdash \varphi(c_1, \ldots, c_k) \quad \textit{if and only if} \quad T \vdash (\forall\, x_1, \ldots, x_k)\varphi(x_1, \ldots, x_k)$$

*Demonstration.* Assume that $T \vdash \varphi(c_1, \ldots, c_k)$. Let's realize that $T$ is a theory in the language $L$ not containing the constants $c_1, \ldots, c_k$, so that the theory $T$ "cannot know anything about them". Therefore, everything what can be proved in $T$ about these constants can be proved about conveniently chosen distinct variables $x_1, \ldots, x_k$ not occurring in the original proof — it suffices to replace every occurrence of the symbol symbol $c_i$ by the variable $x_i$. Then $T \vdash \varphi(x_1, \ldots, x_k)$ and the needed conclusion follows by the Rule of Generalization. The reversed implication is trivial.

It is clear that the above theorem remains true also in case when the constants $c_1, \ldots, c_k$ belong to $L$ but they do not occur in any of the specific axioms of $T$.

When developing and building a mathematical theory we seldom keep its language fixed for all the time. Just the opposite, we often define new notions, corresponding to some operations, relations or distinguished elements, and introduce new symbols for them. These new symbols, as a rule, denote important or frequently occurring concepts, abbreviate otherwise cumbersome formulations and that way contribute to transparency and intelligibility of the theory. Even in our course we already several times did so in the part devoted to various examples of first order theories, without paying special attention to this point. In particular, we extended the language of Group Theory consisting of a single binary operation symbol $\cdot$ by the constant symbol $e$ for the unit element and the unary operation symbol $^{-1}$ for taking inverses, we also extended both the language of the Theory of Boolean Algebras as well as of the language of Peano Arithmetic by the order relation symbol $\leq$ , etc. Now, we will treat this situation in general.

Let $L = (F, C, R, \nu)$ and $L' = (F', C', R', \nu')$ be two first order languages. We say the language $L'$ is an *extension* of the language $L$ if $F \subseteq F'$, $C \subseteq C'$, $R \subseteq R'$ and for each operational or relational symbol $s \in F \cup R$ we have $\nu'(s) = \nu(s)$, i.e., the arities of the symbol $s$ in $L$ and $L'$ coincide. Is it the case, we write $L \subseteq L'$ or $L' \supseteq L$. Then any first order theory $T$ in the language $L$ can be considered as a theory in the language $L' \supseteq L$. The other way round, from any $L'$-structure $\mathcal{A} = (A;\, I)$ one can obtain an $L$-structure $\mathcal{A} \restriction L = (A;\, I \restriction L)$, called the *restriction* of $\mathcal{A}$ to $L$, by leaving its base set $A$ and the interpretations $s^I$ of all the symbols of $L$ unchanged and omitting the interpretations of the remaining symbols of $L'$. We are particularly interested in the case when the new symbols extending $L$ are introduced by means of definitions by $L$-formulas in $T$.

The *unique existence quantification* $(\exists!\, x)\varphi$ is introduced as the abbreviation for $(\exists\, x)\big(\varphi \wedge (\forall\, y)(\varphi(y/x) \Rightarrow y = x)\big)$ where $y$ is any variable not occurring in $\varphi$.

Let $T$ be a theory in a first order language $L$. Dealing with constant, functional and relational symbols we distinguish three possibilities:

(a) Let $\varphi(x)$ be an $L$-formula such that $T \vdash (\exists!\, x)\varphi(x)$. We extend the language $L$ by a new constant symbol $d$ not occurring in $L$ and the theory $T$ by the axiom

$$x = d \;\Leftrightarrow\; \varphi(x)$$

(b) Let $\psi(x_1, \ldots, x_n, y)$ be an $L$-formula such that $T \vdash (\forall\, x_1, \ldots, x_n)(\exists!\, y)\psi(\vec{x}, y)$. We extend the language $L$ by a new $n$-ary functional symbol $g$ not occurring in $L$ and the theory $T$ by the axiom

$$y = g(x_1, \ldots, x_n) \;\Leftrightarrow\; \psi(x_1, \ldots, x_n, y)$$

(c) Let $\rho(x_1, \ldots, x_n)$ be any $L$-formula. We extend the language $L$ by a new $n$-ary relational symbol $q$ not occurring in $L$ and the theory $T$ by the axiom

$$q(x_1, \ldots, x_n) \;\Leftrightarrow\; \rho(x_1, \ldots, x_n)$$

In any of the above cases we say that the constant symbol $d$ or the functional symbol $g$ or the relational symbol $q$, respectively, were introduced by the corresponding definition in $T$. The reader should realize that, regarding constant symbols as nullary functional symbols, (a) can be considered as a special case of (b).

We say that a theory $T'$ in a first order language $L'$ is an *extension* of the theory $T$ in the first order language $L$ *by definitions* if $L'$ is an extension of $L$ by finitely many specific symbols and the specific axioms of $T'$ are obtained extending $T$ by consecutive introduction of the new symbols of $L'$ by definitions. Thus introducing a new symbol at some step we can use not just the means of the original language $L$ in its definition but also the previously introduced symbols. Now it is clear that every model $\mathcal{A} = (A;\, I)$ of the theory $T$ in the language $L$ has a unique extension to a model $\mathcal{A}' = (A;\, I')$ of $T'$ in the language $L'$. It is obtained by repeated interpretation of each newly introduced symbol in $\mathcal{A}'$ using its defining formula in the language $L$ extended by the previously introduced symbols.

Although the new theory $T'$ enables to express several concepts in a more concise and readable way, concerning statements in the original language $L$, it cannot prove more then the original theory $T$.

A theory $T'$ in a first order language $L'$ extending a theory $T$ in a first order language $L \subseteq L'$ is called a *conservative extension* of $T$ if for any closed $L$-formula $\varphi$ we have

$$T' \vdash \varphi \quad \text{if and only if} \quad T \vdash \varphi$$

Obviously, every conservative extension of a consistent theory is itself consistent.

**Theorem on Extension by Definitions.** *Assume that the theory $T'$ in a first order language $L'$ is obtained as an extension by definitions of a theory $T$ in a first order language $L \subseteq L'$. Then $T'$ is a conservative extension of $T$.*

In order to demonstrate the Theorem it would be enough to deal with the case when $L'$ and $T'$ are obtained from $L$ and $T$ by introducing a single defined symbol. The idea of the proof is simple: it consists in replacing every instance the defined formula $x = d$, $y = g(x_1, \ldots, x_n)$ or $q(x_1, \ldots, x_n)$, respectively, by an appropriate instance of the corresponding $L$-formula defining it. However, its realization would require to take care of some technical details which we skip as they would not contribute to the reader's understanding the issue.

**Exercise.** Extensions by defined constants, operations or relation are fairly frequent in Set Theory.

(a) Write explicitly the defining formulas for the empty set constant $\emptyset$, the operations of the unordered pair of elements $\{x, y\}$, of the union $\bigcup X$, as well as for all the operations of taking the subset $\{x \in X \colon \varphi(x, \vec{u})\}$ from the Scheme of Comprehension.

(b) Using appropriate instances of the Scheme of Comprehension introduce the binary operations of intersection $X \cap Y = \{u \colon u \in X \wedge u \in Y\}$ and set-theoretical difference $X \smallsetminus Y = \{u \colon u \in X \wedge u \notin Y\}$.

(c) Write the defining formula for the subset relation $X \subseteq Y$ and, using it, introduce the power set operation $\mathcal{P}(X)$.

(d) Using the operations of unordered pair $\{x, y\}$ and union $\bigcup X$, introduce the binary operation of union $X \cup Y = \{u \colon u \in X \vee u \in Y\}$.

(e) Using the operation of unordered pair, introduce the operation of ordered pair as $(x, y) = \{\{x\}, \{x, y\}\}$ and prove that

$$(x, y) = (u, v) \;\Leftrightarrow\; x = u \wedge y = v$$

(f) Using the operations of ordered pair $(x, y)$, binary union $X \cup Y$ and power set $\mathcal{P}(X)$, as well as an appropriate instance of the Scheme of Comprehension, introduce and justify the operation of cartesian product

$$X \times Y = \{(x, y) \colon x \in X \wedge y \in Y\}$$

**Gödel's Completeness Theorem**

Assume that $L = (F, C, R, \nu)$ is a first order language containing at least one constant symbol (i.e., $C \neq \emptyset$). Denote by $K$ the set of all constant terms of $L$. Then $K$ becomes the base set of an $L$-structure $\mathcal{K} = (K; \dots)$, obtained by interpreting the specific symbols of $L$ in the following natural way:

(a) for any $n$-ary functional symbol $f \in F$ and constant terms $t_1, \dots, t_n \in K$, $f^{\mathcal{K}}(t_1, \dots, t_n)$ is the constant term $f(t_1, \dots, t_n) \in K$;

(b) for any constant symbol $c \in C$, $c^{\mathcal{K}}$ is the constant term $c \in K$;

(c) for any $n$-ary relational symbol $r \in R$ and constant terms $t_1, \dots, t_n \in K$, we put $(t_1, \dots, t_n) \in r^{\mathcal{K}}$ if and only if $T \vdash r(t_1, \dots, t_n)$.

Additionally we introduce the following binary relation $\sim_T$ on $K$:

$$t_1 \sim_T t_2 \;\Leftrightarrow\; T \vdash t_1 = t_2$$

for $t_1, t_2 \in K$. If there's no danger of confusion, we write just $\sim$ instead of $\sim_T$.

**Exercise.** Using the Axioms of Equality show that

$$\begin{aligned}
t &\sim t \\
t_1 \sim t_2 \;&\Rightarrow\; t_2 \sim t_1 \\
(t_1 \sim t_2 \wedge t_2 \sim t_3) \;&\Rightarrow\; t_1 \sim t_3 \\
(t_1 \sim s_1 \wedge \dots \wedge t_n \sim s_n) \;&\Rightarrow\; f(t_1, \dots, t_n) \sim f(s_1, \dots, s_n) \\
(t_1 \sim s_1 \wedge \dots \wedge t_n \sim s_n) \;&\Rightarrow\; (r(t_1, \dots, t_n) \Leftrightarrow r(s_1, \dots, s_n))
\end{aligned}$$

for any $t, t_1, t_2, t_3 \in K$, all $n$-ary symbols $f \in F$, $r \in R$ and any $t_1, s_1, \dots, t_n, s_n \in K$.

The first three conditions express the fact that $\sim$ is reflexive, symmetric and transitive, i.e., it is an *equivalence relation* on the set $K$. For each $t \in K$ we denote by

$$\widetilde{t} = \{s \in K : s \sim t\}$$

the set of all constant terms equivalent with $t$, i.e., all such $s \in K$ for which the equality $s = t$ can be proved in $T$. We always have $t \in \widetilde{t}$, $\widetilde{t} = \widetilde{s}$ if and only if $t \sim s$, and $\widetilde{t} \cap \widetilde{s} = \emptyset$ if $t \not\sim s$. Thus we can form the *quotient set*

$$K/{\sim} = \left\{\widetilde{t} : t \in K\right\},$$

i.e., the *partition* of $K$ into blocks of pairwise equivalent elements. Alternatively, $K/{\sim}$ can be viewed as the result of identifying or merging in a single element all pairwise equivalent elements of $K$, i.e., considering the equivalence relation $\sim$ as a "new equality" on $K$. The last two *compatibility conditions* express the fact that both the operations and the relations in $\mathcal{K}$ preserve the equivalence relation, i.e., the "new equality" $\sim$.

The quotient $M = K/{\sim}$ becomes the base set of an $L$-structure $\mathcal{M} = (M; \dots)$, again, obtained by interpreting the specific symbols of $L$ in the following natural way:

(a) for any $n$-ary functional symbol $f \in F$ and equivalence blocks $\widetilde{t_1}, \dots, \widetilde{t_n} \in M$, $f^{\mathcal{M}}(\widetilde{t_1}, \dots, \widetilde{t_n})$ is the block $f(t_1, \dots, t_n)^{\sim} \in M$ of the constant term $f(t_1, \dots, t_n) \in K$;

(b) for any constant symbol $c \in C$, $c^{\mathcal{M}}$ is the block $\widetilde{c} \in M$ of the constant term $c \in K$;

(c) for any $n$-ary relational symbol $r \in R$ and equivalence blocks $\widetilde{t_1}, \dots, \widetilde{t_n} \in M$, we put $(\widetilde{t_1}, \dots, \widetilde{t_n}) \in r^{\mathcal{M}}$ if and only if $T \vdash r(t_1, \dots, t_n)$.

The compatibility conditions for $\sim$ guarantee that the above definitions of the interpretations $f^{\mathcal{M}}$, $r^{\mathcal{M}}$ of the operation and predicate symbols, respectively, are correct, i.e., they do not depend on the particular representatives of the equivalence blocks $\widetilde{t_i}$.

In order to stress the role of the theory $T$ in the construction of the structure $\mathcal{M}$, we denote it by $\mathcal{M}(T) = \mathcal{M} = (M; \dots)$ and call it the *canonical structure* of the theory $T$.

**Example.** The constant terms in the language of Peano Arithmetic PA are composed of the constant symbols 0 and 1 by means of the operations of addition and multiplication. For instance, 1, 0+1, 1+0, 1·1, 0+(1·1) are five different constant terms, however, they all denote the same natural number 1, and, at the same time, the equality between any pair of them is provable in PA. In other words, $1 \sim_{\mathrm{PA}} 0{+}1 \sim_{\mathrm{PA}} 1{+}0 \sim_{\mathrm{PA}} 1{\cdot}1 \sim_{\mathrm{PA}} 0{+}(1{\cdot}1)$. In fact, there are infinitely many constant terms $t$ in the language of PA such that $1 \sim_{\mathrm{PA}} t$. Similarly, the natural number 2 denotes the equivalence block of the constant term $1+1$ or of any constant term provably equivalent to it, etc. The reader should realize that the canonical structure $\mathcal{M}(\mathrm{PA})$ of the theory PA coincides with its standard model $(\mathbb{N}; +, \cdot, 0, 1)$.

It would be nice if we could guarantee that $\mathcal{M}(T) \vDash T$, i.e., that the canonical structure $\mathcal{M}(T)$ is a model of $T$ for any consistent first order theory $T$ (in a language $L$ containing at least one constant symbol). Unfortunately, this is not always the case. Nevertheless, we can prove that $\mathcal{M}(T) \vDash T$ for theories satisfying a couple of conditions, to be formulated below.

The first of these conditions is the completeness of the theory $T$, similarly as in Propositional Calculus. The second condition has no propositional analogue. Given an $L$-formula $\varphi(x)$ (with a single free variable $x$), a constant $L$-term $t$ is called a *witness* of the sentence $(\exists\, x)\varphi(x)$ in the theory $T$ if

$$T \vdash (\exists\, x)\varphi(x) \Rightarrow \varphi(t)$$

(Notice that the substitution of a constant term for any variable is always admissible.) Since we always have $T \vdash \varphi(t) \Rightarrow (\exists\, x)\varphi(x)$, $t$ is a witness of $(\exists\, x)\varphi(x)$ if and only if

$$T \vdash (\exists\, x)\varphi(x) \Leftrightarrow \varphi(t)$$

Is it the case, then we have $T \vdash (\exists\, x)\varphi(x)$ if and only if $T \vdash \varphi(t)$.

A theory $T$ in a first order language $L$ (containing at least one constant symbol) is called a *Henkin theory* if every $L$-sentence of the form $(\exists\, x)\varphi(x)$ has a witness in $T$.

**Proposition.** *Let $T$ be a complete Henkin theory in a first order language $L$ (containing at least one constant symbol). Then $\mathcal{M}(T) \models T$, in other words, the canonical structure $\mathcal{M}(T)$ of the theory $T$ is a model of $T$.*

*Demonstration.* We will show that, for any closed $L$-formula $\varphi$, we have

$$(*) \qquad\qquad T \vdash \varphi \qquad \text{if and only if} \qquad \mathcal{M}(T) \models \varphi$$

This already implies the needed conclusion $\mathcal{M}(T) \models T$. We will proceed by induction on the complexity of $\varphi$.

Every closed atomic $L$-formula $\varphi$ has the form $t = s$ or $r(t_1, \dots, t_n)$ where $t$, $s$ and $t_1$, ... , $t_n$ are constant terms and $r$ is an $n$-ary relational symbol. Thus for atomic sentences (*) is true according to the definition of the structure $\mathcal{M}(T)$.

Now, it is enough to perform the induction steps for the logical connectives $\neg$ and $\wedge$, and the existential quantifier $\exists$.

Assuming (*) for $\varphi$, we'll verify it for $\neg\varphi$ by showing that the conditions $T \vdash \neg\varphi$ and $\mathcal{M}(T) \models \neg\varphi$ are equivalent. $T \vdash \neg\varphi$ implies $T \nvdash \varphi$ since $T$ is consistent, the reversed implication follows from the completeness of $T$. Thus the conditions $T \vdash \neg\varphi$ and $T \nvdash \varphi$ are equivalent. However, $T \nvdash \varphi$ is equivalent to $\mathcal{M}(T) \not\models \varphi$ by the inductive assumption, and that is equivalent to $\mathcal{M}(T) \models \neg\varphi$.

Assuming (*) for both $\varphi$ and $\psi$, we'll verify it for $\varphi \wedge \psi$. Obviously, the following conditions are equivalent: $T \vdash \varphi \wedge \psi$; $T \vdash \varphi$ and $T \vdash \psi$; $\mathcal{M}(T) \models \varphi$ and $\mathcal{M}(T) \models \psi$; $\mathcal{M}(T) \models \varphi \wedge \psi$ (the inductive assumption is needed to ensure the equivalence of the second and the third condition).

Finally, assuming (*) for all the sentences $\varphi(t)$ where $t$ is a constant term, we will verify it for the sentence $(\exists\, x)\varphi(x)$. Since $T$ is a Henkin theory, the sentence $(\exists\, x)\varphi(x)$ has some witness $t$ in $T$, hence the condition $T \vdash (\exists\, x)\varphi(x)$ is equivalent to the existence of some constant term $t$ such that $T \vdash \varphi(t)$. By the inductive assumption, this is equivalent to the existence of some constant term $t$ such that $\mathcal{M}(T) \models \varphi(t)$, i.e., $\mathcal{M}(T) \models \varphi\big(\widetilde{t}\,\big)$. Since $\mathcal{M}(T) = (M; \dots)$ and $M = K/{\sim}$ consists entirely of elements of the form $\widetilde{t}$ where $t$ is a constant term, the last condition is equivalent to $\mathcal{M}(T) \models (\exists\, x)\varphi$.

**Exercise.** Assume that $S$ is a contradictory theory in a first order language $L$. Describe its canonical structure $\mathcal{M}(S)$ and realize that it is not a model of $S$. Find complete (hence consistent) theory $T$ in $L$ such that $\mathcal{M}(S) = \mathcal{M}(T) \models T$.

Using the *Axiom of Choice* (one of the higher axioms of Set Theory) it is possible to prove the following theorem. The interested reader will find its proof in the Appendix. Dealing with a fixed first order language $L$, a *new symbol* (no matter whether a constant, functional or relational one) always means a specific symbol not occurring in $L$.

**Theorem on Complete Henkin Extensions.** *Let $T$ be a consistent theory in a first order language $L = (F, C, R, \nu)$. Then there is an extension of $L$ by a set $D$ of new constant symbols to some first order language $L_D = (F, C \cup D, R, \nu)$ and an extension of $T$ to a complete Henkin theory $\widehat{T} \supseteq T$ in the language $L_D$.*

We will use the last Theorem in the demonstration of the following result, which is an alternative version of the Completeness Theorem.

**Gödel's Completeness Theorem.** *Every consistent first order theory $T$ has some model $\mathcal{A} \models T$.*

The reader should realize that also the other way round, if a first order theory has some model then it must be consistent, in other words, a contradictory first order theory cannot have any model. (This is the alternative version of the Soundness Theorem.)

*Demonstration.* Let $T$ be a consistent theory in a first order language $L$, the first order language $L_D$ be an extension of $L$ by certain set $D$ of new constant symbols, and $\widehat{T} \supseteq T$ be a theory in $L_D$ forming a complete Henkin extension of $T$. According to the last Proposition, the canonical structure $\mathcal{M}(\widehat{T})$ of the theory $\widehat{T}$ is a model of $\widehat{T}$, i.e., $\mathcal{M}(\widehat{T}) \models \widehat{T}$. Since $T \subseteq \widehat{T}$, we have $\mathcal{M}(\widehat{T}) \models T$, hence $\mathcal{M}(\widehat{T})$ is a model of $T$, as well.

Those who feel puzzled by the fact that $T$ is a theory in the language $L$, while $\mathcal{M}(\widehat{T})$ is an $L_D$-structure, can form the restriction $\mathcal{M}(\widehat{T}) \restriction L$ of the $L_D$-structure $\mathcal{M}(\widehat{T})$ to the language $L$. Then $\mathcal{A} = \mathcal{M}(\widehat{T}) \restriction L$ is already an $L$-structure and, obviously, $\mathcal{A} \models T$.

Finally we can prove the original form of the Completeness Theorem. We state it in the form comprising the Soundness Theorem, as well.

**Completeness Theorem.** *Let $T$ be a theory in a first order language $L$. Then, for every $L$-formula $\psi$, $T \models \psi$ if and only if $T \vdash \psi$.*

*Demonstration.* If $T \vdash \psi$ then $T \models \psi$ by the Soundness Theorem. To show the converse, assume that $T \models \psi$, nevertheless $T \nvdash \psi$. Without loss of generality we can assume that $\psi$ is closed. (Otherwise, we can replace $\psi$ by the sentence $(\forall x_1, \ldots, x_n)\psi$, which we denote by $\overline{\psi}$, where $x_1, \ldots, x_n$ are all the variables occurring freely in $\psi$. Then we have $T \vdash \psi$ if and only if $T \vdash \overline{\psi}$, and $T \models \psi$ if and only if $T \models \overline{\psi}$.) As $\psi$ is closed, from $T \nvdash \psi$ it follows that the theory $T \cup \{\neg\psi\}$ is consistent by the Theorem on Proof by Contradiction. Then, according to Gödel's Completeness Theorem, $T \cup \{\neg\psi\}$ has some model $\mathcal{A}$. Then $\mathcal{A} \models T$ is a model of the theory $T$ such that $\mathcal{A} \models \neg\psi$. However, since $T \models \psi$, we have $\mathcal{B} \models \psi$ for every model $\mathcal{B}$ of $T$; in particular, $\mathcal{A} \models \psi$. This contradiction proves that $T \vdash \psi$.

## The Compactness Theorem

Once we have established Gödel's Completeness Theorem, the first order version of the Compactness Theorem can be demonstrated as its corollary in essentially the same way as its Propostional Calculus version. We leave it to the reader as an exercise.

**Compactness Theorem.** *Let $T$ be a theory in a first order language $L$. Then $T$ has some model if and only if every finite subtheory $T_0$ of $T$ has some model.*

However, unless its Predicate Calculus version, the first order version of the Compactness Theorem has several important consequences. We confine ourselves to just some few examples.

To start with, the reader should realize the following immediate consequence of the Compactness Theorem.

**Corollary.** *Let $T$, $S$ be two theories in a first order language $L$. Then the theory $T \cup S$ has some model if and only if, for every finite subtheory $U$ of $S$, the theory $T \cup U$ has some model.*

A first order theory is said to have *arbitrarily big finite models* if for every natural number $n \geq 1$ there is a model $\mathcal{A} = (A; \dots)$ of the theory $T$ such that $|A| \geq n$.

**Theorem.** *Let $T$ be a theory in a first order language $L$. If $T$ has arbitrarily big finite models, then $T$ has some infinite model, as well.*

*Demonstration.* For every $n \geq 2$ we denote by $\sigma_n$ the following sentence in the language of pure equality:

$$(\exists\, x_1, \dots, x_n)\Big(\bigwedge_{1 \leq i < j \leq n} x_i \neq x_j\Big)$$

Then, for any $L$-structure $\mathcal{A} = (A; \dots)$, we have $\mathcal{A} \vDash \sigma_n$ if and only if $|A| \geq n$.

For each $n \geq 2$ we denote by $S_n$ the first order theory with axioms $\sigma_2, \dots, \sigma_n$ and by $S = \{\sigma_n : 2 \leq n \in \mathbb{N}\}$ the theory formed by all the axioms $\sigma_n$. Obviously, a first order structure $\mathcal{A}$ is infinite if and only if $\mathcal{A} \vDash S$.

Assume that $T$ has arbitrarily big models. This is to say that each of the theories $T \cup S_n$, where $n \geq 2$, has some model. Then, however, for every finite subtheory $U \subseteq S$, there is some $n \geq 2$ such that $U \subseteq S_n$. Since any model of the theory $T \cup S_n$ is a model of $T \cup U$, the theory $T \cup U$ has some model. By the Compactness Theorem the theory $T \cup S$ has some model $\mathcal{A}$, as well. This $\mathcal{A}$ is an infinite model of $T$.

**Exercise.** (a) Show that if a first order theory $T$ in the language of the Theory of Unitary Rings has as its models unitary rings of arbitrarily big finite characteristic then it has as a model also a unitary ring of characteristic $\infty$.

(b) Show that the characteristic of a field is either a prime or $\infty$ and prove that if a first order theory $T$ in the language of the Theory of Unitary Rings has as its models fields of arbitrarily big prime characteristic then it has as a model also a field of characteristic $\infty$.

(c) Show the following *Robinson's Principle*: If a sentence $\varphi$ in the language of the Theory of Unitary Rings is satisfied in every field of characteristic $\infty$ then there is a prime number $p$ such that $\varphi$ is satisfied in every field of prime characteristic $q \geq p$.

Another striking consequence of the Compactness Theorem is the existence of nonstandard models of Peano Arithmetic.

**Theorem.** *Peano Arithmetic has some nonstandard models.*

*Demonstration.* Let us extend the language of PA by a new constant symbol $q$. Let $\chi_n$ denote the formula $q \neq n$ in the extended language (recall that, for every natural number $n$, we denote by $n$ also the constant term $(\dots(0+1)+\dots+1)+1$ in the language of PA obtained by adding 1 repeatedly $n$ times to 0).

We introduce the theories $S = \{\chi_n : n \in \mathbb{N}\}$ and $S_n = \{\chi_0, \chi_1, \dots, \chi_n\}$ for each $n \in \mathbb{N}$. Interpreting $q^{\mathcal{M}_n}$ as the natural number $n+1$, we obtain the model

$$\mathcal{M}_n = (\mathbb{N}; +, \cdot, 0, 1, n+1) \vDash \mathrm{PA} \cup S_n$$

of the theory PA $\cup S_n$. By the Compactness Theorem it follows, that also the theory PA $\cup S$ has some model $\mathcal{M} = (M; +, \cdot, 0, 1, q)$. In this model, the interpretation $q^{\mathcal{M}} \in M$ of the symbol $q$ differs from all the constant terms $n \in \mathbb{N}$, thus $(M; +, \cdot, 0, 1)$ is a nonstandard model of PA.

Intuitively, $(M; +, \cdot, 0, 1)$ can be viewed as a number system extending the standard natural number system $(\mathbb{N}; +, \cdot, 0, 1)$ by some ideal "infinite" natural numbers, with one of them represented by (the interpretation of) the constant symbol $q$. Then of course, $q - 1$, $q + 1$, $2q$, $q^2$, etc., represent different infinite elements of $M$. Moreover, if $p \in M$ is any infinite element then so is $p - 1$, since if $p - 1$ were finite then $p = (p - 1) + 1$ would be finite, too. Thus the nonempty set of all infinite elements of $M$ has no least element, seemingly contradicting the Well Ordering Principle implied by the Scheme of Induction of PA. Nonetheless, this paradox has a simple resolution: the sets of finite and infinite elements in $M$, respectively, are not first order expressible. This means that there are no formulas $\varphi(x, u_1, \ldots, u_n)$, $\psi(x, u_1, \ldots, u_n)$ in the language of PA and no (finite or infinite) elements $p_1, q_1, \ldots, p_n, q_n \in M$ such that

$$\{a \in M : a \text{ is finite}\} = \{a \in M : \mathcal{M} \vDash \varphi(a, p_1, \ldots, p_n)\}$$
$$\{a \in M : a \text{ is infinite}\} = \{a \in M : \mathcal{M} \vDash \psi(a, q_1, \ldots, q_n)\}$$

Similar accounts show that there are also nonstandard number systems $({}^*\mathbb{R}; +, \cdot, 0, 1)$, extending the standard number system $(\mathbb{R}; +, \cdot, 0, 1)$ of all real numbers and satisfying all the axioms of the Theory of Real Closed Fields, having the same first order properties as $(\mathbb{R}; +, \cdot, 0, 1)$. Such number systems contain, besides standard reals, also infinite (infinitely big) and infinitesimal (infinitely small) number quantities. Using them, it is possible (among other things) to develop the infinitesimal (i.e., the differential and the integral) calculus in an intuitively appealing way, close to its historically original form, in the spirit of Newton, Leibniz, Euler and others, and that way to rehabilitate and justify the approach abandoned during the $19^{\text{th}}$ century in favor of the techniques of limits and the $\varepsilon\delta$-analysis.

## Cardinality of Models and Skolem's Paradox

A detailed inspection of the proof of Gödel's Completeness Theorem (both in Section ... as well as in the Appendix) would show that the construed model satisfies an additional cardinality specification.

The *cardinality of a first order language* $L = (F, C, R, \nu)$ is defined as

$$\|L\| = |\operatorname{Form}(L)| = \max(|F|, |C|, |R|, \aleph_0)$$

A *first order language* $L$ is called *countable* if $\|L\| = \aleph_0$. Obviously, any first order language with just finitely many specific symbols is countable.

It is clear that the set $K$ of all constant terms of any first order language $L$ has the cardinality $|K| \leq |\operatorname{Term}(L)|$. The base set $M$ of the canonical structure $\mathcal{M}(T) =$

$(M; \dots)$ of any theory $T$ in the language $L$ is a quotient $M = K/{\sim_T}$, therefore $|M| \le |K|$. Thus the canonical structure $\mathcal{M}(T) = (M; \dots)$ of any theory in a first order order language $L$ has the cardinality

$$|M| \le |K| \le |\operatorname{Term}(L)| \le |\operatorname{Form}(L)| = \|L\|$$

Similarly, the set $D$ of new constant symbols, added to the language $L$ for the sake of construction of the complete Henkin extension $T^+$ of $T$, has the same cardinality $\|L\|$. Thus the new language $L^+$ has the same cardinality as the original language $L$. Putting things together, we see that the canonical structure $\mathcal{M}(T^+) = (M; \dots)$ still has the cardinality $|M| \le \|L\|$. As a consequence, we obtain the following strengthening of Gödel's Completeness Theorem.

**Theorem.** *Let $T$ be a consistent theory in a first order language $L$ of cardinality $\|L\| = \alpha$. Then $T$ has a model $\mathcal{M} = (M; \dots)$ of cardinality $|M| \le \alpha$.*

**Corollary 1.** *Every consistent first order theory in a countable language has a countable model, i.e., a model $\mathcal{M} = (M; \dots)$ of cardinality at most $\aleph_0$.*

Realizing that the usual language of Set Theory has a single specific symbol, namely the binary relational symbol $\in$ for the membership relation, we readily obtain:

**Corollary 2.** *Any of the set theories* ZF, ZFC *(if it is consistent) has a countable model $\mathcal{M} = \big(M; \in^{\mathcal{M}}\big)$.*

Moreover, by *Mostowski Collapse Theorem*, we can arrange that $a \subseteq M$ for $a \in M$, and $a \in^{\mathcal{M}} b$ if and only if $a \in b$ for all $a, b \in M$. Then, according to the Soundness Theorem, everything that can be proved in Set Theory must be satisfied in $\mathcal{M}$. In particular, there are sets $\mathbb{N}^{\mathcal{M}}$, $\mathbb{R}^{\mathcal{M}} \in M$ playing in $\mathcal{M}$ the role of the set of all natural numbers and of the set of all real numbers, respectively. However, since $\mathbb{N}^{\mathcal{M}} \subseteq M$, $\mathbb{R}^{\mathcal{M}} \subseteq M$, both the sets $\mathbb{N}^{\mathcal{M}}$, $\mathbb{R}^{\mathcal{M}}$ are countable, hence (as it is clear that none of them can be finite) there is a bijective mapping $f\colon \mathbb{N}^{\mathcal{M}} \to \mathbb{R}^{\mathcal{M}}$. On the other hand, Cantor's Theorem *"the set $\mathbb{R}$ of all real numbers is uncountable"*, which is provable in Set Theory, must be true in $\mathcal{M}$, as well. This sounds like a contradiction.

This paradox was discovered by the Norwegian mathematician Thoralf Skolem in 1922. However, Skolem derived it from the *Löwenheim-Skolem Downward Theorem* (which we will deal with later on) and not from Gödel's Completeness Theorem (though it was known to him well before Gödel proved and published it in 1930, but he neither proved it nor formulated it explicitly). Skolem's Paradox is not a contradiction proving the inconsistency of Set Theory. It can be resolved in the following way: The bijection $f\colon \mathbb{N}^{\mathcal{M}} \to \mathbb{R}^{\mathcal{M}}$ does not belong to the model $\mathcal{M}$; in fact there is no function $f \in M$ establishing a bijective correspondence $f\colon \mathbb{N}^{\mathcal{M}} \to \mathbb{R}^{\mathcal{M}}$. Thus Cantor's Theorem still holds in $\mathcal{M}$. Informally, the set $\mathbb{R}^{\mathcal{M}}$ is uncountable just from the internal point of view (i.e., as a set belonging to the model $\mathcal{M}$), while from the external point of view it is still countable. Nevertheless, Skolem's Paradox indicates that the notions like countability or uncountability, similarly as several other set-theoretical concepts concerning infinite cardinal numbers, lack an absolute character.

## Appendix

## Proof of the Theorem on Complete Henkin Extensions

Let $(A; \leq)$ be a partially ordered set. An element $m \in A$ is called *maximal* if there is no element $a \in A$ such that $m < a$. A subset $C \subseteq A$ is called a *chain* if $a \leq b$ or $b \leq a$ holds for any $a, b \in C$, i.e., if $C$ is totally ordered by the relation $\leq$.

Any subset $\mathcal{T} \subseteq \mathcal{P}(X)$ of the powerset of any set $X$ will be referred to a sa a *system of subsets* of $X$ and automatically regarded as a partially ordered set $(\mathcal{T}; \subseteq)$ with the relation of set-theoretical inclusion. We say that a system $\mathcal{T} \subseteq \mathcal{P}(X)$ of subsets of a set $X$ has *finite character* if for any $T \subseteq X$ we have $T \in \mathcal{T}$ if and only if $F \in \mathcal{T}$ for any finite set $F \subseteq T$. We say that a system $\mathcal{T} \subseteq \mathcal{P}(X)$ is *inductive* if for any chain $\mathcal{C} \subseteq \mathcal{T}$ also its union $\bigcup \mathcal{C}$ belongs to $\mathcal{T}$.

We record without proof the following two consequences of the Axiom of Choice with the remark that anyone of them is in fact equivalent to it.

**Teichmüller-Tukey Lemma.** *Let $X$ be any set and $\mathcal{T} \subseteq \mathcal{P}(X)$ be a system of finite character of subsets of $X$. Then for every $T \in \mathcal{T}$ there exists a maximal element $M \in \mathcal{T}$ such that $T \subseteq M$.*

**Zorn-Kuratowski Lemma.** *Let $X$ be any set and $\mathcal{T} \subseteq \mathcal{P}(X)$ be an inductive system of subsets of $X$. Then for every $T \in \mathcal{T}$ there exists a maximal element $M \in \mathcal{T}$ such that $T \subseteq M$.*

Let us take for $X$ the set $\Phi = \mathrm{Form}(L)$ of all formulas of some first order language $L$ and denote by $\mathcal{T} \subseteq \mathcal{P}(\Phi)$ the system of all consistent first order theories in $L$. As already noticed in the proof of the Compactness Theorem, a theory is consistent if and only if every its finite subtheory $T_0 \subseteq T$ is consistent. In other words, the system $\mathcal{T} \subseteq \mathcal{P}(\Phi)$ of all consistent theories in $L$ has finite character.

The other way round, given a chain $\mathcal{C}$ of consistent theories in $L$, it can easily be seen that its union $U = \bigcup \mathcal{C}$ is a consistent theory, again. Indeed, if $U$ were contradictory then there would be some sentence $\varphi$ and a proof $\varphi_0, \varphi_1, \dots, \varphi_m$ of $\varphi$ in $U$, as well as a proof $\psi_0, \psi_1, \dots, \psi_n$ of its negation $\neg\varphi$ in $U$. Let $\chi_1, \dots, \chi_k$ be all the axioms of $U$ occurring in any of these proofs. Then there are theories $T_1, \dots, T_k \in \mathcal{C}$ such that $\chi_i \in T_i$ for each $i = 1, \dots, k$. Since $\mathcal{C}$ is a chain, there is some $p$ such that $1 \leq p \leq k$ and $T_i \subseteq T_p$ for any $i = 1, \dots, k$. Then $\{\chi_1, \dots, \chi_k\} \subseteq T_p$, hence both $\varphi_0, \varphi_1, \dots, \varphi_m$ and $\psi_0, \psi_1, \dots, \psi_n$ are already proofs in $T_p$, so that $T_p$ is contradictory. However, this is impossible, as $T_p \in \mathcal{C}$ and all the elements of $\mathcal{C}$ are consistent theories. Summing up, the system of all consistent theories $\mathcal{T} \subseteq \mathcal{P}(\Phi)$ is inductive.

Now, both the Teichmüller-Tukey Lemma as well as the Zorn-Kuratowski Lemma imply that every consistent theory $T$ in a first order language $L$ can be extended to a maximal consistent theory $M \supseteq T$ in $L$.

**Exercise.** Let $T$ be a theory in a first order language $L$. We denote by

$$T^{\vdash} = \{\varphi \in \mathrm{Form}(L) \colon T \vdash \varphi\}$$

the set of all $L$-formulas provable in $T$. Show that $T$ is complete if and only if $T^{\vdash}$ is a maximal consistent theory.

The above Exercise has furnished us with the last piece of knowledge needed in order to establish the following result.

**Lindenbaum's Theorem.** *Every consistent theory $T$ in a first order language $L$ can be extended to a complete theory $T^+ \supseteq T$ in $L$.*

**Remark.** Although (some equivalent alternative formulations of) the Axiom of Choice played a crucial role in the proof of Lindenbaum's Theorem, it is known that this theorem is weaker than AC, in the sense that AC cannot be proved in Zermelo-Fraenkel Set Theory assuming Lindenbaum's Theorem.

Next we show a result on Henkin extensions of consistent theories.

**Theorem on Conservative Henkin Extensions.** *Let $T$ be a consistent theory in a first order language $L = (F, C, R, \nu)$. Then there is an extension of $L$ by a set $D$ of new constant symbols to some first order language $L_D = (F, C \cup D, R, \nu)$ and a conservative extension of $T$ to a Henkin theory $T_H \supseteq T$ in the language $L_D$.*

*Demonstration.* Our aim is to equip every sentence $(\exists\, x)\varphi(x)$ with a new witnessing constant $d_\varphi$ and to extend $T$ by the corresponding witnessing axiom $(\exists\, x)\varphi(x) \Rightarrow \varphi(d_\varphi)$. However, doing so for all $L$-formulas $\varphi(x)$, the language $L$ is likely extended and new sentences $(\exists\, x)\varphi(x)$ calling for their own witnessing constants arise. That's why we have to iterate the extension procedure recursively.

Let us denote by $L_0 = L$ the original first order language and by $\Phi_0$ the set of all $L_0$-formulas with a single free variable $x$. For every formula $\varphi \in \Phi_0$ we introduce a new constant symbol $d_\varphi$ in such way that for different formulas $\varphi, \psi \in \Phi_0$ the symbols $d_\varphi$, $d_\psi$ are distinct, as well. We denote by $D_0 = \{d_\varphi \colon \varphi \in \Phi_0\}$ the set of all these constants, by $L_1$ the extension of the language $L_0$ by the set $D_0$ of the new constants and by

$$W_0 = \{(\exists\, x)\varphi(x) \Rightarrow \varphi(d_\varphi) \colon \varphi \in \Phi_0\}$$

the set of all the witnessing axioms for the sentences $(\exists\, x)\varphi(x)$ where $\varphi \in \Phi_0$.

Assuming that the set of formulas $\Phi_n$, the set of constant symbols $D_n = \{d_\varphi \colon \varphi \in \Phi_n\}$, the language $L_{n+1}$ and the set of witnessing axioms $W_n = \{(\exists\, x)\varphi(x) \Rightarrow \varphi(d_\varphi) \colon \varphi \in \Phi_n\}$ are already defined, we denote by $\Phi_{n+1}$ the set of all formulas of the language $L_{n+1}$ with a single free variable $x$ not belonging to the union $\bigcup_{k=0}^{n} \Phi_k$. For every $\varphi \in \Phi_{n+1}$ we introduce a new (i.e., not occurring in the language $L_{n+1}$) constant symbol $d_\varphi$, with distinct symbols $d_\varphi$, $d_\psi$ corresponding to different formulas $\varphi$, $\psi$. Next we denote by $D_{n+1} = \{d_\varphi \colon \varphi \in \Phi_{n+1}\}$ the set of all the recently added constants, by $L_{n+2}$ the extension of the language $L_{n+1}$ by the set $D_{n+1}$ of these constants and by

$$W_{n+1} = \{(\exists\, x)\varphi(x) \Rightarrow \varphi(d_\varphi) \colon \varphi \in \Phi_{n+1}\}$$

the set of all witnessing axioms for the sentences $(\exists\, x)\varphi(x)$ where $\varphi \in \Phi_{n+1}$.

Finally we put $D = \bigcup_{n\in\mathbb{N}} D_n$, $W = \bigcup_{n\in\mathbb{N}} W_n$ and denote by $L_D$ the extension of the language $L$ by the set of the new constant symbols $D$. We claim that $T_H = T \cup W$ is a Henkin theory in the language $L_H$ and a conservative extension of $T$.

It can be easily seen that every $L_D$-sentence of the form $(\exists\, x)\varphi(x)$ has a witness in the theory $T_H$. Since $\varphi$ contains just finitely many constant symbols from $D$ (if any), there is the smallest $n \in \mathbb{N}$ such that $\varphi$ contains no symbol from $D_m$ for all $m \geq n$. Then $\varphi \in \Phi_n$ and the sentence $(\exists\, x)\varphi(x) \Rightarrow \varphi(d_\varphi)$ belongs to $W_n$ hence to $T_H$. Thus the constant symbol $d_\varphi \in D$ is a witness of the sentence $(\exists\, x)\varphi(x)$ in $T_H$.

In order to show that $T_H$ is a conservative extension of $T$ it is enough to verify that any $L$-sentence $\psi$ provable in $T_H$ is provable already in $T$. If $T_H \vdash \psi$ then there are finitely many witnessing axioms $\theta_i$ of the form $(\exists\, x)\varphi_i(x) \Rightarrow \varphi_i(d_i)$ with $1 \leq i \leq k$, where we write $d_i$ instead of $d_{\varphi_i}$, such that $T \cup \{\theta_1, \ldots, \theta_k\} \vdash \psi$. If $k = 0$ then already $T \vdash \psi$ and we are done; thus we can assume that $k \geq 1$. Then it is enough to show that the number $k$ of witnessing axioms can be anytime reduced by 1.

There's again the smallest $n$ such that none of the formulas $\varphi_1, \ldots, \varphi_k$ contain any constant symbol from $D_m$ for all $m \geq n$. Then $\varphi_j \in \Phi_n$ for some $j \in \{1, \ldots, k\}$ and none of the witnessing sentences $\theta_i$ for $i \neq j$ contains the symbol $d_j$. For brevity's sake we denote $\varphi_j$ by $\varphi$, $d_j$ by $d$ and $\Theta = \{\theta_i \colon 1 \leq i \leq k,\ i \neq j\}$. As a consequence of the Deduction Theorem, we have

$$T \cup \Theta \vdash \big((\exists\, x)\varphi(x) \Rightarrow \varphi(d)\big) \Rightarrow \psi$$

Now the reader is asked to realize that, for any propositional forms $A$, $B$, $C$, both the propositional forms

$$((A \Rightarrow B) \Rightarrow C) \Rightarrow (\neg A \Rightarrow C) \qquad \text{and} \qquad ((A \Rightarrow B) \Rightarrow C) \Rightarrow (B \Rightarrow C)$$

are tautologies, therefore, by the Completeness Theorem for Propositional Calculus, they are provable just from the propositional logical axioms. In particular, both the formulas

$$\Big(\big((\exists\, x)\varphi(x) \Rightarrow \varphi(d)\big) \Rightarrow \psi\Big) \Rightarrow \big(\neg(\exists\, x)\varphi(x) \Rightarrow \psi\big)$$
$$\Big(\big((\exists\, x)\varphi(x) \Rightarrow \varphi(d)\big) \Rightarrow \psi\Big) \Rightarrow (\varphi(d) \Rightarrow \psi)$$

are provable just from the propositional axioms of Predicate Calculus. Then, by Modus Ponens, we have both

$$T \cup \Theta \vdash \neg(\exists\, x)\varphi(x) \Rightarrow \psi \qquad \text{and} \qquad T \cup \Theta \vdash \varphi(d) \Rightarrow \psi$$

Since the symbol $d$ doesn't occur in any of the specific axioms of the theory $T \cup \Theta$, applying the Theorem on Constants to the latter relation we obtain

$$T \cup \Theta \vdash (\forall\, x)(\varphi(x) \Rightarrow \psi)$$

Since $\psi$ is closed, the variable $x$ is not free in $\psi$, thus, according to (d) and the third equivalence in (b) of Exercise on the prenex normal form,

$$T \cup \Theta \vdash (\exists\, x)\varphi(x) \Rightarrow \psi$$

Since for any propositional forms $A$, $B$ the propositional form

$$(A \Rightarrow B) \Rightarrow ((\neg A \Rightarrow B) \Rightarrow B)$$

is a tautology, hence provable just from the logical axioms(cf. Exercise... (g)), we have

$$T \cup \Theta \vdash ((\exists\, x)\varphi(x) \Rightarrow \psi) \Rightarrow ((\neg(\exists\, x)\varphi(x) \Rightarrow \psi) \Rightarrow \psi)$$

Applying Modus Ponens twice we get $T \cup \Theta \vdash \psi$. Since the set $\Theta$ consists of $k-1$ witnessing axioms, only, we are done.

**Exercise.** Show directly, i.e. without referring to the last Theorem, that the Henkin extension $T_H$ of the consistent theory $T$ constructed in its demonstration is consistent. (It can be done in a similar but simpler way than was the demonstration of conservativeness of the extension $T_H$.)

Combining the two recently established theorems, we can finally prove the announced result.

**Theorem on Complete Henkin Extensions.** *Let $T$ be a consistent theory in a first order language $L = (F, C, R, \nu)$. Then there is an extension of $L$ to some first order language $L_D = (F, C \cup D, R, \nu)$ by a set $D$ of new constant symbols and an extension of $T$ to a complete Henkin theory $\widehat{T} \supseteq T$ in the language $L_D$.*

*Demonstration.* Let $T_H$ be the conservative Henkin extension of the theory $T$ in the language $L_D = (F, C \cup D, R, \nu)$ construed as above. Then, by Lindenbaum's Theorem, there is a complete theory $\widehat{T} = T_H^+ \supseteq T_H$ in the same language $L_H$. Obviously, $\widehat{T}$ is a Henkin theory, as well.

## Selected Bibliography


B. Balcar, P. Štěpánek, *Teorie množin*, Academia, Praha 2000 (2. vyd.).

L. Bukovský, *Množiny a všeličo okolo nich*, Alfa, Bratislava 1985.

C. C. Chang, H. J. Keisler, *Model Theory*, Studies in Logic and the Foundations of Mathematics, Elsevier, North Holland, Amsterdam 1990 (3rd ed.).

H.-D. Ebbinghaus, J. Flum, *Finite Model Theory*, Perspectives in Mathematical Logic, Springer, Berlin-Heidelberg-New York 1999 (2nd ed.).

H. B. Enderton, *A Mathematical Introduction to Logic*, Hartcourt/Academic Press, San Diego-New York-Boston-London-Toronto-Sydney-Tokyo 2001 (2nd ed.).

T. Franzen, *Gödel's Theorem: An Incomplete Guide to Its Use and Abuse*, A K Peters, Wellesley, Mass. 2005.

M. Hills, F. Loeser, *A First Journey through Logic*, Student Mathematical Library vol. 89, Amer. Math. Soc. 2019.

W. Hodges, *Model Theory*, Cambridge University Press, Cambridge-New York 1993.

W. Hodges, *A Shorter Model Theory*, Cambridge University Press, Cambridge-New York 1997.

J. Kolář, O. Štepánková, M. Chytil, *Logika, algebry a grafy* (Kap. 2. Výroková logika, Kap. 3. Predikátová logika), SNTL, Praha 1989

V. Kolman, V. Punčochář, *Formy jazyka: Úvod do logiky a její filosofie*, Filosofia, Praha 2015.

Yu. I. Manin, *A Course in Mathematical Logic for Mathematicians*, Graduate Texts in Mathematics, Springer, New York-Dordrecht-Heidelberg-London 2010 (2nd ed. with collaboration by B. Zilber).

D. Marker, *Model Theory, An Introduction*, Graduate Texts in Mathematics, Springer, New York-London 2002.

E. Mendelson, *Introduction to Mathematical Logic*, D. Van Nostrand, Princeton, New Jersey 1964.

G. Restal, *Logic, An Introduction*, Routlege, Taylor & Francis Group, London-New York 2005.

J. R. Shoenfield, *Mathematical Logic*, Addison-Wesley, Reading, Mass. 1967, reprinted by Assoc. Symb. Logic, Urbana, Illinois and A K Peters, Natick, Mass. 2001.

P. Smith, *An Introduction to Gödel's Theorems*, Cambridge University Press, Cambridge-New York 2013.

R. M. Smullyan, *Gödel's Incompleteness Theorems*, Oxford University Press, Oxford-New York 1992.

A. Sochor, *Klasická matematická logika*, Univerzita Karlova, Nakladatelství Karolinum, Praha 2001.

A. Sochor, *Logika pro všechny ochotné myslet*, Univerzita Karlova, Praha 2011.

P. Štěpánek, *Matematická logika*, skriptá MFF UK, SPN, Praha 1982.

V. Švejdar, *Logika – neúplnost, složitost, nutnost*, Academia, Praha 2002.

Hao Wng, *A Logical Journey: From Gödel to Philosophy*, MIT Press, Cambridge, Mass., London, England 1996.

P. Zlatoš, *Ani matematika si nemôže byť istá sama sebou: Úvahy o množinách, nekonečne, paradoxoch a Gödelových vetách*, IRIS, Bratislava 1995.

P. Zlatoš, *O dobrom usporiadaní a axióme výberu*, Obzory matematiky, fyziky a informatiky, **58** (1999), 1–9 (1. časť), **1** (2000), 14–24 (2. časť).