2. THE LOWER PREDICATE CALCULUS WITH IDENTITY, LPCI

Contents

§1. Examples of structures. The Lower Predicate (or First Order) Calculus is interpreted in *mathematical structures* like the following.

1A. Graphs. A *graph* (for these notes) is a pair

$$\mathbf{G} = (G, E)$$

where $G \neq \emptyset$ is a non-empty set (the *nodes* or *vertices*) and $E \subseteq G^2$ is a binary relation on G, (the *edges*) such that for all nodes x, $\neg E(x, x)$; $\mathbf{G}$ is *symmetric* or *unordered* if

$$E(x, y) \Longrightarrow E(y, x).$$

A *path* in a symmetric graph $\mathbf{G} = (G, E)$ is a sequence of nodes

$$(x_0, x_1, \dots, x_n)$$

such that there is an edge joining each x_i with x_{i+1}, i.e.,

$$E(x_0, x_1), E(x_1, x_2), \dots, E(x_{n-1}, x_n);$$

a path *joins* its first vertex x_0 with its last x_n.

The *distance* between two distinct nodes x, y in a symmetric graph $\mathbf{G}$ is the length (number of edges, n above) of the shortest path joining them, if some path joins them,

$$d(x, y) = \min\{n \mid \text{there exists a path } (x_0, \dots, x_n) \text{ with } x_0 = x, x_n = y\}.$$

By convention, $d(x, x) = 0$ and if no path joins x to y we set $d(x, y) = \infty$.

A symmetric graph is *connected* if any two points in it are joined by a path.

1B. Partial and linear orderings. A *partial ordering* or *partially ordered set* or *poset* is a pair

$$\mathbf{P} = (P, \leq),$$

where P is a non-empty set and $\leq$ is a binary relation on P satisfying the following conditions:

(a) For all $x \in P$, $x \leq x$ (reflexivity).
(b) For all $x, y, z \in P$, if $x \leq y$ and $y \leq z$, then $x \leq z$ (transitivity).
(c) For all $x, y \in P$, if $x \leq y$ and $y \leq x$, then $x = y$ (antisymmetry).

A *total* or *linear ordering* is a partial ordering in which every two elements are comparable, i.e., such that

(d) For all $x, y \in P$, either $x \le y$ or $y \le x$.

1C. Fields. A *field* is a tuple

$$\mathbf{K} = (K, 0, 1, +, \cdot)$$

where K is a set, $0, 1 \in K$, $+$ and $\cdot$ are binary operations on K and the following *field axioms* are true.

(1) $(K, 0, +)$ is a *commutative group*, i.e.,

(a) For all x, $x + 0 = x$.
(b) For all x, y, z, $x + (y + z) = (x + y) + z$.
(c) For all x, y, $x + y = y + x$.
(d) For every x, there exists some y such that $x + y = 0$.

(2) $1 \neq 0$ and for all x, $x \cdot 0 = 0$, $x \cdot 1 = x$.

(3) The structure $(K \setminus \{0\}, 1, \cdot)$ is a commutative group, and in particular

$$x, y \neq 0 \Longrightarrow x \cdot y \neq 0.$$

Together with (2), this means that for all x, y in K,

$$x \cdot y = 0 \iff x = 0 \text{ or } y = 0.$$

(4) *The distributive law*: For all x, y, z, $x \cdot (y + z) = x \cdot y + x \cdot z$.

We will also use the following notations for the standard number fields of the *rational*, the *real* and the *complex numbers*:

$$\mathbf{Q} = (\mathbb{Q}, 0, 1, +, \cdot), \quad \mathbf{R} = (\mathbb{R}, 0, 1, +, \cdot), \quad \mathbf{C} = (\mathbb{C}, 0, 1, +, \cdot).$$

1D. The natural numbers. The structure of *arithmetic* or the *natural numbers* is the tuple

$$\mathbf{N} = (\mathbb{N}, 0, S, +, \cdot)$$

where $\mathbb{N} = \{0, 1, 2, \dots\}$ is the set of (non-negative) integers and S, $+$, $\cdot$ are the operations of successor, addition and multiplication on $\mathbb{N}$. The structure $\mathbf{N}$ has the following characteristic properties:

(1) The successor function S is an injection, i.e.,

$$S(x) = S(y) \Longrightarrow x = y,$$

and 0 is not a successor, i.e., for all x, $S(x) \neq 0$.

(2) *Induction Principle*: for every set of numbers $X \subseteq \mathbb{N}$, if $0 \in X$ and for every x, $x \in X \Longrightarrow S(x) \in X$, then $X = \mathbb{N}$.

(3) For all x, y, $x + 0 = x$ and $x + S(y) = S(x + y)$.

(4) For all x, y, $x \cdot 0 = 0$ and $x \cdot S(y) = x \cdot y + x$.

These properties (or sometimes just the first two of them) are called the *Peano Axioms* for the natural numbers.

Some of the most significant applications of logic that we will prove are about $\mathbf{N}$, the natural domain of interpretation of the *theory of numbers*.

§2. Syntax of LPCI. The name LPCI abbreviates *Lower Predicate Calculus with Identity*. It is actually a family of languages $\mathsf{LPCI}(\tau)$, one for each *vocabulary* or *signature* τ, where the signature provides names for the distinguished elements, relations and functions of the structures we want to talk about.

LPCI is also known as *First Order Logic with Identity*, or *Elementary Logic with Identity*.

2A. Vocabularies. A *vocabulary* or *signature* is a quadruple

$$\tau = (\mathrm{Const}, \mathrm{Rel}, \mathrm{Funct}, \mathrm{arity}),$$

such that the sets of (individual) *constant symbols* Const, *relation symbols* Rel, and *function symbols* Funct are disjoint and

$$\mathrm{arity} : \mathrm{Rel} \cup \mathrm{Funct} \to \{1, 2, \dots\}.$$

A relation or function symbol P is n-ary if $\mathrm{arity}(P) = n$.

Most often we will assume that τ is **finite**, i.e., the sets Const, Rel and Funct are finite, but it is useful to allow the possibility that they are infinite; and we should also keep in mind that any one—or all—of these sets may be empty.

We often specify signatures and languages by enumerating their symbols, when they are finitely many and their arities are clear from the context: for example,

- $\mathsf{LPCI}(E)$ is the *language of graphs* (with E binary);
- $\mathsf{LPCI}(\leq)$ is the *language of posets*;
- $\mathsf{LPCI}(0, 1, +, \cdot)$ is the *language of fields*:

and more importantly for what we will do,

- $\mathsf{LPCI}(0, S, +, \cdot)$ is the *language of arithmetic*.

2B. Terms and formulas. The *alphabet* of $\mathsf{LPCI}(\tau)$ comprises the symbols in the vocabulary τ and the following, additional symbols which are common to all these languages:

(a) The *logical symbols* $\neg \ \wedge \ \vee \ \rightarrow \ \forall \ \exists \ =$
(b) The *punctuation symbols* ( ) ,
(c) The (individual) *variables*: $\mathrm{v}_0, \mathrm{v}_1, \mathrm{v}_2, \dots$

The new logical symbols $\forall$ and $\exists$ are the **quantifiers** and they are read "for all" and "there exists".

As with PL, *words* are finite strings of symbols and $\mathrm{lh}(\alpha)$ is the length of the word α; we use $\equiv$ to denote identity of strings,

$$\alpha \equiv \beta \iff_{\mathrm{df}} \alpha \text{ and } \beta \text{ are the same string;}$$

and we set

$$\alpha \sqsubseteq \beta \iff_{\text{df}} \alpha \text{ is an initial segment of } \beta,$$

so that e.g., $\forall v_0 \sqsubseteq \forall v_0 R(v_0)$. The *concatenation* of two strings $\alpha\beta$ is the string produced by putting them together, with α first, so that $\alpha \sqsubseteq \alpha\beta$.

Terms are defined by the recursion:

(a) Each variable v_i is a term.
(b) Each constant symbol is a term.
(c) If $t_1, \dots, t_n$ are terms and f is an n-ary function symbol, then $f(t_1, \dots, t_n)$ is a term.
(d) No string is a term except by virtue of (a) – (c).

As with the definition of PL-formulas, we abbreviate this by

$$t :\equiv v \mid c \mid f(t_1, \dots, t_n) \tag{2-1}$$

and we interpret it rigorously as in Definition 1A.1 of Part 1.

LPCI-**Formulas** are defined by the recursion:

(a) If s, t are terms, then $s = t$ is a formula.
(b) If $t_1, \dots, t_n$ are terms and R is an n-ary relation symbol, then $R(t_1, \dots, t_n)$ is a formula.
(c) If ϕ, ψ are formulas and v is a variable, then the following strings are formulas:

$$(\neg\phi) \quad (\phi \wedge \psi) \quad (\phi \vee \psi) \quad (\phi \to \psi) \quad \forall v\phi \quad \exists v\phi$$

(d) No string is a formula except by virtue of (a) – (c).

In abbreviated form:

$$\phi :\equiv t_1 = t_2 \mid R(t_1, \dots, t_n) \\ \mid (\neg\phi) \mid (\phi \wedge \psi) \mid (\phi \vee \psi) \mid (\phi \to \psi) \mid \forall v\phi \mid \exists v\phi$$

Terms and formulas are collectively called (well formed) **expressions**.

2B.1. **Proposition** (Unique readability for terms)**.** *Each term t satisfies exactly one of the following three conditions.*

(a) *$t \equiv v$ for a uniquely determined variable v.*
(b) *$t \equiv c$ for a uniquely determined constant c.*
(c) *$t \equiv f(t_1, \dots, t_n)$ for a uniquely determined function symbol f and uniquely determined terms $t_1, \dots, t_n$.*

2B.2. **Proposition** (Unique readability for formulas)**.** *Each formula χ satisfies exactly one of the following conditions.*

(a) *$\chi \equiv s = t$ for uniquely determined terms s, t.*
(b) *$\chi \equiv R(t_1, \dots, t_n)$ for a uniquely determined relation symbol R and uniquely determined terms $t_1, \dots, t_n$.*
(c) *$\chi \equiv (\neg\phi)$ for a uniquely determined formula ϕ.*
(d) *$\chi \equiv (\phi \wedge \psi)$ for uniquely determined formulas ϕ, ψ.*

(e) $\chi \equiv (\phi \lor \psi)$ *for uniquely determined formulas* ϕ, ψ.
(f) $\chi \equiv (\phi \to \psi)$ *for uniquely determined formulas* ϕ, ψ.
(g) $\chi \equiv \exists v \phi$ *for a uniquely determined variable* v *and a uniquely determined formula* ϕ.
(h) $\chi \equiv \forall v \phi$ *for a uniquely determined variable* v *and a uniquely determined formula* ϕ.

A formula is **prime** if it satisfies (a) or (b) in this lemma and **quantifier-free** if no quantifier occurs in it.

These *Parsing Lemmas* are proved very much like Theorem 1A.3 of Part 1 and they allow us to give definitions by (structural) recursion on terms and formulas.

2B.3. **Abbreviations and misspellings**. As in the propositional calculus, we will almost never spell formulas correctly: we will use infix notation for terms, e.g.,

$$s + t \text{ for } +(s, t)$$

in arithmetic, introduce and use abbreviations, use (meta)variables (names) $x, y, z, u, v, \ldots,$ for the formal variables of the language, skip (or add) parentheses, replace parentheses by brackets or other punctuation marks, and (in general) be satisfied with giving "instructions" for writing out a term or a formula rather than exhibiting the actual formal expression. For example, the following formula in the language of arithmetic says that there are infinitely many prime numbers:

$$(\forall x)(\exists y)\Big[x \leq y \land (\forall u)(\forall v)[(y = u \cdot v) \to (u = 1 \lor v = 1)\Big]$$

where we have used the abbreviations

$$x \leq y :\equiv (\exists z)[x + z = y] \qquad 1 :\equiv S(0)$$

The correctly spelled formula which corresponds to this is quite long (and unreadable).

Two useful logical abbreviations are for the "iff"

$$(\phi \leftrightarrow \psi) :\equiv ((\phi \to \psi) \land (\psi \to \phi))$$

and the quantifier "there exists exactly one x such that ϕ"

$$(\exists! x)\phi :\equiv (\exists z)(\forall x)[\phi \leftrightarrow x = z], \tag{2-2}$$

where $z \not\equiv x$. (Think this through.)

2C. Free and bound occurrences of variables. Every occurrence of a variable in a term is *free*.

The free occurrences of variables in formulas are defined by structural recursion on formulas, as follows.

(a) $\mathrm{FO}(s = t) = \mathrm{FO}(s) \cup \mathrm{FO}(t)$.

(b) $\mathrm{FO}((\neg\phi)) = \mathrm{FO}(\phi)$, $\mathrm{FO}((\phi \wedge \psi)) = \mathrm{FO}(\phi) \cup \mathrm{FO}(\psi)$, and similarly $\mathrm{FO}((\phi \vee \psi)) = \mathrm{FO}((\phi \to \psi)) = \mathrm{FO}(\phi) \cup \mathrm{FO}(\psi)$.

(c) $\mathrm{FO}(\forall v\phi) = \mathrm{FO}(\exists v\phi) = \mathrm{FO}(\phi) \setminus \{v\}$, meaning that we remove from the free occurrences of variables in ϕ all the occurrences of the variable v.

An occurrence of a variable which is not free in an expression α is *bound*.

For example, the free occurrences of variables are underlined in the formula $(\exists \mathrm{v}_2 P(\mathrm{v}_2, \underline{\mathrm{v}_1}, \mathrm{v}_2) \wedge R(\underline{\mathrm{v}_2}))$. Notice that in this example, the variable v_2 has (three) bound occurrences and one free one.

To understand the significance of this distinction between free and bound occurrences of variables, consider the expression

$$(*) \qquad \int_0^a x^2 dx$$

we use in Calculus. There are two variables which occur in it, 'x' (twice) and 'a' (once). Now,

$$\int_0^a x^2 dx = \int_0^a y^2 dy \quad \Big(= \frac{a^3}{3} \Big),$$

i.e., the value of the integral does not change if we replace x by y in the formula for it; while, if $a \neq b$, then

$$\int_0^a x^2 dx = \frac{a^3}{3} \neq \frac{b^3}{3} = \int_0^b x^2 dx.$$

On the account that we are giving, this is because the occurrences of x are *bound* in the formula (∗) for the integral, while that of a is free.

The **free variables** of an expression α are the variables which have at least one free occurrence in α; the **bound variables** of α are those which have at least one bound occurrence in α; and α is **closed** if it has no free variables.

A closed formula is called a **sentence**.

2D. Substitutions. For each expression α, each variable v and each term t, the expression $\alpha\{v :\equiv t\}$ is the result of replacing all free occurrences of v in α by the term t. The simultaneous substitution

$$\alpha\{v_1 :\equiv t_1, \dots, v_n :\equiv t_n\}$$

is defined similarly: we replace simultaneously all the free occurrences of each v_i in α by t_i. If α is an expression, then $\alpha\{v_1 :\equiv t_1, \dots, v_n :\equiv t_n\}$ is also an expression, and of the same kind—term or formula. Note that by Problem x2.6, in general

$$\alpha\{v_1 :\equiv t_1\}\{v_2 :\equiv t_2\} \not\equiv \alpha\{v_1 :\equiv t_1, v_2 :\equiv t_2\}.$$

A term t_i is **free for** v_i in a substitution if no occurrence of a variable in t_i is bound in the result of the substitution. We will only perform *free substitutions* of this kind, sometimes without explicitly stating the hypotheses.

2E. Extended terms and formulas. If α is a term or formula and $v_1, \dots, v_n$ is a list of distinct variables which includes all the free variables of α, we set

$$\alpha(v_1, \dots, v_n) :\equiv (\alpha, (v_1, \dots, v_n)).$$

So $\alpha(v_1, \dots, v_n)$ is a pair of an expression and a sequence of variables.

This is a very useful device: for example, if $\alpha(v_1, \dots, v_n)$ is an extended expression and $t_1, \dots, t_n$ a sequence of terms free for $v_1, \dots, v_n$ in α, we can then set

$$\alpha(t_1, \dots, t_n) :\equiv \alpha\{v_1 :\equiv t_1, \dots, v_n :\equiv t_n\}$$

exhibiting the result of the free substitution operation in a useful way.

2F. Lower predicate calculus without identity, LPC. We will also work with the smaller language LPC, which is obtained by removing the symbol = and the clauses involving it in the definitions. There are no formulas in LPC(τ), unless the signature τ has at least one relation symbol. In stating theorems about LPC we will tacitly assume that the signature has at least one relation symbol, so we have formulas.

§3. Semantics of LPCI. As we have already mentioned, we will interpret the terms and formulas of LPCI(τ) in structures of signature τ, defined as follows:

3A. Structures. A *structure* of signature

$$\tau = (\text{Const}, \text{Rel}, \text{Funct}, \text{arity})$$

(or τ-**structure**) is a pair $\mathbf{A} = (A, I)$, where A is a non-empty set and the *interpretation* I assigns

(a) a member $I(c) = c^{\mathbf{A}}$ of A to each constant symbol c;
(b) an n-ary relation $I(R) = R^{\mathbf{A}} \subseteq A^n$ to each n-ary relation symbol,
and
(c) an n-ary function $I(f) = f^{\mathbf{A}} : A^n \to A$ to each function symbol f of arity n.

The set A is the *universe* of the structure $\mathbf{A}$ and the constants, relations and functions which interpret the symbols of the signature in $\mathbf{A}$ are its *primitives.* We often use the succinct notation

$$\mathbf{A} = (A, \{c^{\mathbf{A}}\}_{c \in \text{Const}}, \{R^{\mathbf{A}}\}_{R \in \text{Rel}}, \{f^{\mathbf{A}}\}_{f \in \text{Funct}}) \tag{3-1}$$

and if τ has only finitely many symbols, we denote $\mathbf{A}$ as a tuple of its universe and its primitives as in the examples of Section 1,

$$\mathbf{G} = (G, E), \quad \mathbf{K} = (K, 0, 1, +, \cdot), \quad \mathbf{N} = (\mathbb{N}, 0, S, +, \cdot).$$

We use the notation in (3-1) in the next, basic definition.

3B. Isomorphisms. An *isomorphism*

$$\sigma : \mathbf{A} \rightarrowtail\!\!\!\!\twoheadrightarrow \mathbf{B}$$

between two τ-structures $\mathbf{A}, \mathbf{B}$ is a bijection $\sigma : A \rightarrowtail\!\!\!\!\twoheadrightarrow B$ such that:

(a) For each constant symbol c, $\sigma(c^{\mathbf{A}}) = c^{\mathbf{B}}$.

(b) For each n-ary relation symbol R and all $x_1, \dots, x_n \in A$,

$$R^{\mathbf{A}}(x_1, \dots, x_n) \iff R^{\mathbf{B}}(\sigma(x_1,), \dots, \sigma(x_n)).$$

(c) For each n-ary function symbol f and all $x_1, \dots, x_n \in A$,

$$\sigma(f^{\mathbf{A}}(x_1, \dots, x_n)) = f^{\mathbf{B}}(\sigma(x_1), \dots, \sigma(x_n)).$$

An isomorphism $\sigma : \mathbf{A} \rightarrowtail\!\!\!\!\twoheadrightarrow \mathbf{A}$ of a structure with itself is called an *automorphism* of $\mathbf{A}$.

Two structures of the same signature are *isomorphic* if there is an isomorphism from one to the other,

(3-2) $$\mathbf{A} \cong \mathbf{B} \iff \text{there exists an isomorphism } \sigma : \mathbf{A} \rightarrowtail\!\!\!\!\twoheadrightarrow \mathbf{B}.$$

3C. Reducts and expansions. If σ and τ are vocabularies and each symbol of σ is a symbol (of the same kind and with the same arity) in τ, we say that σ is a *reduct* of τ and τ is an *expansion* of σ and we write $\sigma \subseteq \tau$. For example, $(0, 1, \cdot) \subseteq (0, 1, +, \cdot)$.

Suppose $\sigma \subseteq \tau$, $\mathbf{A} = (A, I)$ is a σ-structure and $\mathbf{B} = (B, J)$ is a τ-structure. We call $\mathbf{A}$ a *reduct* of $\mathbf{B}$ and $\mathbf{B}$ an *expansion* of $\mathbf{A}$ if $A = B$ and for all symbols $C \in \sigma$, $I(C) = J(C)$. If $\mathbf{B}$ is a given τ-structure and $\sigma \subseteq \tau$, we define *the reduct of* $\mathbf{B}$ *to* σ by deleting from $\mathbf{B}$ the objects assigned to the symbols not in σ, formally $\mathbf{B} \restriction \sigma = (B, J \restriction \sigma)$; for example,

$$\mathbf{R} \restriction (0, +) = (\mathbb{R}, 0, +)$$

is the additive group of real numbers.

Conversely, if $\tau \subseteq \sigma$, we can define expansions of $\mathbf{B}$ by assigning interpretations to the symbols in σ which are not in τ. The notation for this is

$$(\mathbf{B}, K) =_{\text{df}} \text{the expansion of } \mathbf{B} \text{ by } K.$$

For example, the ordered real field

$$\mathbf{R}_o = (\mathbf{R}, \leq) = (\mathbb{R}, 0, 1, +, \cdot, \leq)$$

is the expansion of the real field $\mathbf{R}$ by the ordering relation and

$$(\mathbf{N}, \exp) = (\mathbb{N}, 0, S, +, \cdot, \exp)$$

is the expansion of the structure of arithmetic by the exponential function, $\exp(t, x) = x^t$.

3D. Substructures and extensions. Suppose $\mathbf{A} = (A, I)$, $\mathbf{B} = (B, J)$ are τ-structures. We call $\mathbf{A}$ a *substructure* of $\mathbf{B}$ and write $\mathbf{A} \subseteq \mathbf{B}$ if the following conditions hold.

(a) $A \subseteq B$.
(b) For every constant symbol c of τ, $c^{\mathbf{B}} = c^{\mathbf{A}} \in A$.
(c) For every n-ary relation symbol R and all $x_1 \dots, x_n \in A$,

$$R^{\mathbf{B}}(x_1 \dots, x_n) \iff R^{\mathbf{A}}(x_1 \dots, x_n).$$

(d) For every n-ary function symbol f and all $x_1 \dots, x_n \in A$,

$$f^{\mathbf{B}}(x_1 \dots, x_n) = f^{\mathbf{A}}(x_1 \dots, x_n) \in A.$$

For example, $\mathbf{Q} \subseteq \mathbf{R}$, i.e., the field of rational numbers is a substructure of the field of real numbers.

The difference between reducts and substructures is important.

3E. Denotations of terms and formulas. An **assignment** into a structure $\mathbf{A}$ is any association of objects in A with the variables, i.e., any function π : Variables $\to A$.

The *value* or *denotation* of a term for an assignment π in a structure $\mathbf{A}$ is defined by structural recursion on the terms as follows:

(a) $\text{value}(v, \pi) =_{\text{df}} \pi(v)$,
(b) $\text{value}(c, \pi) =_{\text{df}} c^{\mathbf{A}}$,
(c) $\text{value}(f(t_1, \dots, t_n), \pi) =_{\text{df}} f^{\mathbf{A}}(\text{value}(t_1, \pi), \dots, \text{value}(t_n, \pi))$.

In the same way, we define the *truth value* or *denotation* (1 or 0) of a formula for an assignment π in a structure $\mathbf{A}$ by the following structural recursion:

(a) $\text{value}(s = t, \pi) =_{\text{df}} \begin{cases} 1, & \text{if value}(s, \pi) = \text{value}(t, \pi), \\ 0, & \text{otherwise.} \end{cases}$
(b) $\text{value}(R(t_1, \dots, t_n), \pi) =_{\text{df}} \begin{cases} 1, & \text{if } R^{\mathbf{A}}(\text{value}(t_1, \pi), \dots, \text{value}(t_n, \pi)), \\ 0, & \text{otherwise.} \end{cases}$
(c) $\text{value}(\neg\phi, \pi) =_{\text{df}} 1 - \text{value}(\phi, \pi) = \begin{cases} 0, & \text{if value}(\phi, \pi) = 1, \\ 1, & \text{otherwise.} \end{cases}$
(d) $\text{value}(\phi \wedge \psi, \pi) = \min(\text{value}(\phi, \pi), \text{value}(\psi, \pi))$. For $\vee$ we take the maximum and for implication we use

$$\text{value}(\phi \to \psi, \pi) =_{\text{df}} \text{value}(\neg(\phi) \vee \psi).$$

(e) $\text{value}(\exists v(\phi), \pi) =_{\text{df}} \max\{\text{value}(\phi, \rho) \mid \text{ for all } v' \not\equiv v, \rho(v') = \pi(v')\}$.
(f) $\text{value}(\forall v(\phi), \pi) =_{\text{df}} \min\{\text{value}(\phi, \rho) \mid \text{ for all } v' \not\equiv v, \rho(v') = \pi(v')\}$.

When we need to exhibit the dependence of denotations on $\mathbf{A}$ we write

$$\text{value}^{\mathbf{A}}(\alpha, \pi) = \text{value}(\alpha, \pi).$$

3F. Satisfaction and the Tarski conditions. For a given τ-structure $\mathbf{A}$, an assignment π into $\mathbf{A}$ and a τ-formula ϕ, we put

$$\text{(3-3)}\quad \mathbf{A}, \pi \models \phi \iff_{\text{df}} \text{value}^{\mathbf{A}}(\phi, \pi) = 1$$
$$\iff \text{the assignment } \pi \text{ satisfies the formula } \phi \text{ in the structure } \mathbf{A}.$$

This ternary *satisfaction relation* on structures, assignments and formulas is the most fundamental notion of the semantics of LPCI. To formulate its characteristic properties we need the following notion:

Updates. For a variable v and any $x \in A$, the *update* $\pi\{v := x\}$ is the assignment which agrees with π on all variables except v to which it assigns x,

$$\pi\{v := x\}(\mathrm{v}_i) = \begin{cases} x, & \text{if } \mathrm{v}_i \equiv v, \\ \pi(\mathrm{v}_i), & \text{otherwise;} \end{cases}$$

more generally, for a sequence $\vec{v} \equiv v_1, \ldots, v_n$ of distinct variables and a corresponding sequence $\vec{x} = x_1, \ldots, x_n$ of elements of A, the *simultaneous update*

$$\pi\{\vec{v} := \vec{x}\} = \pi\{v_1 := x_1, \ldots, v_n := x_n\}$$

is defined by

$$\pi\{\vec{v} := \vec{x}\}(\mathrm{v}_i) = \begin{cases} x_i, & \text{if } \mathrm{v}_i \equiv v_i \text{ for some } i = 1, \ldots, n, \\ \pi(\mathrm{v}_i), & \text{otherwise.} \end{cases}$$

3F.1. **Theorem** (The Tarski conditions). *For every τ-structure $\mathbf{A}$, every assignment π into $\mathbf{A}$, and all formulas ϕ, ψ:*

$$\begin{aligned}
\mathbf{A}, \pi \models s = t &\iff \text{value}^{\mathbf{A}}(t, \pi) = \text{value}^{\mathbf{A}}(s, \pi) \\
\mathbf{A}, \pi \models \neg\phi &\iff \mathbf{A}, \pi \not\models \phi \\
\mathbf{A}, \pi \models \phi \wedge \psi &\iff \mathbf{A}, \pi \models \phi \textit{ and } \mathbf{A}, \pi \models \psi \\
\mathbf{A}, \pi \models \phi \vee \psi &\iff \mathbf{A}, \pi \models \phi \textit{ or } \mathbf{A}, \pi \models \psi \\
\mathbf{A}, \pi \models \phi \rightarrow \psi &\iff \textit{either } \mathbf{A}, \pi \not\models \phi \textit{ or } \mathbf{A}, \pi \models \psi \\
\mathbf{A}, \pi \models \exists v \phi &\iff \textit{there exists an } x \in A \textit{ such that } \mathbf{A}, \pi\{v := x\} \models \phi \\
\mathbf{A}, \pi \models \forall v \phi &\iff \textit{for all } x \in A, \mathbf{A}, \pi\{v := x\} \models \phi
\end{aligned}$$

PROOF is by structural induction on formulas, cf. Problem x2.10. ⊣

3G. Compositionality. Denotations of terms and formulas are defined in such a way that *the value of an expression is a function of the values of its subexpressions.* This is generally referred to as the *Compositionality Principle* for denotations, and it is the key to their mathematical analysis. The next theorem gives two, different precise versions of it.

3G.1. **Theorem.** (1) *If the σ-structure* $\mathbf{A}$ *is a reduct of the τ-structure* $\mathbf{B}$ *where $\sigma \subseteq \tau$ and ϕ is a formula of* $\mathsf{LPCI}(\sigma)$*, then for every expression α and every assignment π,*

$$\mathrm{value}^{\mathbf{A}}(\alpha, \pi) = \mathrm{value}^{\mathbf{B}}(\alpha, \pi).$$

(2) *If π, ρ are two assignments into the same structure* $\mathbf{A}$ *and for every variable v which occurs free in an expression α, $\pi(v) = \rho(v)$, then*

$$\mathrm{value}^{\mathbf{A}}(\alpha, \pi) = \mathrm{value}^{\mathbf{A}}(\alpha, \rho).$$

Proof of both claims is by structural induction on α. ⊣

In particular, if χ is a sentence, then $\mathrm{value}(\mathbf{A}, \chi, \pi)$ is independent of the assignment π and we set

$$\text{(3-4)}\qquad \begin{aligned} \mathbf{A} \models \chi &\iff_{\text{df}} \text{for some assignment } \pi, \mathbf{A}, \pi \models \chi \\ &\iff \text{for every assignment } \pi, \mathbf{A}, \pi \models \chi \\ &\iff_{\text{df}} \chi \text{ is } \textbf{true} \text{ in } \mathbf{A}. \end{aligned}$$

3H. Denotations of extended expressions. If $t(\vec{v})$ is an extended term by Definition 2E (so that the variables which occur in t are all included in the list $\vec{v} = (v_1, \dots, v_n)$) and $\vec{x} = (x_1, \dots, x_n)$ is an arbitrary n-tuple in A, we set

$$t^{\mathbf{A}}[\vec{x}] =_{\text{df}} \mathrm{value}^{\mathbf{A}}(t, \pi\{\vec{v} := \vec{x}\}) \quad \text{(for any assignment } \pi);$$

and if $\chi(v_1, \dots, v_n)$ is an extended formula and $\vec{x} = (x_1, \dots, x_n)$ is an arbitrary n-tuple in A, we set

$$\chi^{\mathbf{A}}[\vec{x}] \iff_{\text{df}} \mathbf{A}, \pi\{\vec{v} := \vec{x}\} \models \chi,$$

noticing that by the definition of updates and Theorem 3G.1, these functions and relations on A do not depend on which π we use to define them.

3I. First-order expressibility. A proposition Φ of ordinary (mathematical) English about a τ-structure $\mathbf{A}$ is *expressed* (or *formalized*) by a sentence ϕ of $\mathsf{LPCI}(\tau)$ if Φ and ϕ "mean" the same thing, for example,

$$(\forall x)[x + 0 = x] \text{ means "every number added to 0 yields itself"}.$$

It is clear that we cannot make this "expressing" precise unless we first define *meaning* rigorously for both natural language and LPCI. On the other hand, we have a good, intuitive understanding of this notion of "expressibility" which is important for applications: roughly speaking, ϕ expresses Φ if we can construct the first from the second by straightforward translation, more-or-less word for word, "and", "but" and "also" going to $\wedge$, "all", "each" and "any" going to $\forall$, etc. For example, "every number is either odd or even" is formalized in the language of arithmetic by something like

$$(\forall x)[\phi(x) \vee \psi(x)],$$

where $\phi(x)$ and $\psi(x)$ can be constructed to express the properties of being odd or even.

The trick to dealing with this vague notion of "formalization" is to use instead the following, related notion which is rigorous and robust.

3J. Elementary relations and functions. A relation $R \subseteq A^n$ on the universe of a structure $\mathbf{A}$ is *first-order definable* or *elementary* on $\mathbf{A}$, if there is an extended formula $\chi(v_1, \dots, v_n)$ such that

$$R(\vec{x}) \iff \chi^{\mathbf{A}}[\vec{x}];$$

and a function $g : A^n \to A$ on the universe of a structure $\mathbf{A}$ is *first-order definable* or *elementary* on $\mathbf{A}$ if its **graph**

$$G_g(\vec{x}, w) \iff g(\vec{x}) = w$$

is elementary on $\mathbf{A}$, i.e., if there is an extended formula $\chi(\vec{v}, u)$ such that

$$g(\vec{x}) = w \iff \chi^{\mathbf{A}}[\vec{x}, w].$$

3J.1. **Theorem.** *The collection $\mathcal{E}(\mathbf{A})$ of $\mathbf{A}$-elementary functions and relations on the universe of a structure*

$$\mathbf{A} = (A, \{c^{\mathbf{A}}\}_{c \in \mathrm{Const}}, \{R^{\mathbf{A}}\}_{R \in \mathrm{Rel}}, \{f^{\mathbf{A}}\}_{f \in \mathrm{Funct}})$$

has the following properties:

(1) *Each primitive relation $R^{\mathbf{A}}$ and the* (binary) *identity relation $x = y$ on A are $\mathbf{A}$-elementary.*

(2) *For each constant symbol c and each n, the n-ary constant function*

$$g(\vec{x}) = c^{\mathbf{A}}$$

is $\mathbf{A}$-elementary; each primitive function $f^{\mathbf{A}}$ is $\mathbf{A}$-elementary; and every projection function

$$P_i^n(x_1, \dots, x_n) = x_i \quad (1 \le i \le n)$$

is $\mathbf{A}$-elementary.

(3) *$\mathcal{E}(\mathbf{A})$ is closed under substitutions of $\mathbf{A}$-elementary functions: i.e., if $h(u_1, \dots, u_m)$ is an m-ary $\mathbf{A}$-elementary function and $g_1(\vec{x}), \dots, g_m(\vec{x})$ are n-ary, $\mathbf{A}$-elementary, then the function*

$$f(\vec{x}) = h(g_1(\vec{x}), \dots, g_m(\vec{x}))$$

is $\mathbf{A}$-elementary; and if $P(u_1, \dots, u_m)$ is an m-ary $\mathbf{A}$-elementary relation, then the n-ary relation

$$Q(\vec{x}) \iff P(g_1(\vec{x}), \dots, g_m(\vec{x}))$$

is $\mathbf{A}$-elementary.

(4) $\mathcal{E}(\mathbf{A})$ *is closed under the propositional operations: i.e., if* $P_1(\vec{x})$ *and* $P_2(\vec{x})$ *are* $\mathbf{A}$*-elementary,* n*-ary relations, then so are the following relations:*

$$\begin{aligned} Q_1(\vec{x}) &\iff \neg P_1(\vec{x}), \\ Q_2(\vec{x}) &\iff P_1(\vec{x}) \wedge P_2(\vec{x}), \\ Q_3(\vec{x}) &\iff P_1(\vec{x}) \vee P_2(\vec{x}), \\ Q_4(\vec{x}) &\iff P_1(\vec{x}) \to P_2(\vec{x}). \end{aligned}$$

(5) $\mathcal{E}(\mathbf{A})$ *is closed under quantification on* A*, i.e., if* $P(\vec{x}, y)$ *is* $\mathbf{A}$*-elementary, then so are the relations*

$$\begin{aligned} Q_1(\vec{x}) &\iff (\exists y)P(\vec{x}, y), \\ Q_2(\vec{x}) &\iff (\forall y)P(\vec{x}, y). \end{aligned}$$

Moreover: $\mathcal{E}(\mathbf{A})$ *is the smallest collection of functions and relations* (of all arities) *on* A *which satisfies* (1) – (5).

Proof. To show that $\mathcal{E}(\mathbf{A})$ has these properties, we need to construct lots of formulas and appeal repeatedly to the definition of $\mathbf{A}$-elementary functions and relations; this is tedious, but not difficult.

For the second ("moreover") claim, we first make it precise by replacing $\mathcal{E}(\mathbf{A})$ by $\mathcal{F}$ throughout (1) – (5), and (temporarily) calling a class $\mathcal{F}$ of functions and relations *good* if it satisfies all these conditions—so what has already been shown is that $\mathcal{E}(\mathbf{A})$ is good. The additional claim is that *every good* $\mathcal{F}$ *contains all* $\mathbf{A}$*-elementary functions and relations*, and it is verified by structural induction on the extended formula $\chi(\vec{v})$ which defines a given, $\mathbf{A}$-elementary relation—after showing, easily, that the graph of every function defined by a term is in $\mathcal{F}$. ⊣

The theorem suggests that $\mathcal{E}(\mathbf{A})$ is a very rich class of relations and functions. As it turns out, this is true for "complex" structures like $\mathbf{N}$ but not true for "simple" structures, like the usual ordering $(\mathbb{Q}, \leq)$ on the rational numbers, and understanding this phenomenon is an important problem in **Model Theory**, one of the main parts of our subject. We will consider it briefly, mostly by examples, in the next two sections, and then again in Part 3 of these class notes.

§4. Arithmetical relations and functions. The elementary relations and functions of the standard structure $\mathbf{N}$ of arithmetic are called *arithmetical*.

Is the exponential function

$$\exp(t, x) = x^t \quad (x, t \in \mathbb{N})$$

arithmetical? Not obviously—but it is, as a corollary of a basic result about *definition by recursion in* $\mathbf{N}$ which we will prove in this section and which has many important applications.

We will need to appeal to several simple but not trivial facts from number theory and we will outline proofs for most of these, but it is not surprizing that we need them: the key Theorem 4A.1 below is an important, basic fact about the natural numbers and so its truth ultimately depends on what the natural numbers are.

4A. Primitive recursion. A function $f : \mathbb{N} \to \mathbb{N}$ is defined by *primitive recursion* from $w_0 \in \mathbb{N}$ and $h : \mathbb{N}^2 \to \mathbb{N}$ if it satisfies the following two equations, for all $t \in \mathbb{N}$:

$$f(0) = w_0, \quad f(t+1) = h(f(t), t). \tag{4-1}$$

In greater generality, a function $f : \mathbb{N}^{n+1} \to \mathbb{N}$ is defined by *primitive recursion* from $g : \mathbb{N}^n \to \mathbb{N}$ and $h : \mathbb{N}^{n+2} \to \mathbb{N}$ if it satisfies the following two equations, for all $t \in \mathbb{N}, x \in \mathbb{N}^n$:

$$f(0, x) = g(x), \quad f(t+1, x) = h(f(t, x), t, x). \tag{4-2}$$

For example, if we set

$$f(0, x) = x, \quad f(t+1, x) = S(f(t, x)) \quad (t, x \in \mathbb{N}),$$

then (easily, by induction on t)

$$f(t, x) = t + x,$$

and so addition is defined by primitive recursion from the two, simpler functions

$$g(x) = x, \quad h(w, t, x) = S(w),$$

i.e., (essentially) the identity and the successor. Similarly, if we set

$$f(0, x) = 0, \quad f(t+1, x) = f(t, x) + x,$$

then, easily, $f(t, x) = t \cdot x$, and so multiplication is defined by primitive recursion from the functions

$$g(x) = 0, \quad h(w, t, x) = w + x,$$

i.e., essentially the constant 0 and addition. More significantly (for our purposes here),

$$\exp(0, x) = x^0 = 1, \quad \exp(t+1, x) = x^{t+1} = x^t \cdot x = \exp(t, x) \cdot x,$$

so that exponentiation is defined by primitive recursion from the functions

$$g(x) = 1, \quad h(w, t, x) = w \cdot x,$$

i.e., (essentially) the constant 1 and multiplication. (This definition of exponentiation gives $0^0 = 1$, which is not a useful convention in calculus but does not introduce any problems in number theory.)

The main result in this section is the following theorem which, in particular, implies that the exponential function is arithmetical.

4A.1. **Theorem.** *If $f : \mathbb{N}^{n+1} \to \mathbb{N}$ is defined by the primitive recursion* (4-2) *above and g, h are arithmetical, then so is f.*

To prove it, we must reduce the recursive definition of f into an explicit one, and this is done using **Dedekind's analysis of recursion**:

4A.2. **Proposition** (Dedekind)**.** *If $f : \mathbb{N}^{n+1}$ is defined by the primitive recursion in* (4-2)*, then for all $t \in \mathbb{N}, x \in \mathbb{N}^n, w \in \mathbb{N}$,*

(4-3) $f(t,x) = w \iff$ *there exists a sequence* $(w_0, \dots, w_t)$ *such that*

$$\Big(w_0 = g(x) \wedge (\forall s < t)[w_{s+1} = h(w_s, s, x)] \wedge w = w_t\Big).$$

Proof. If $f(t,x) = w$, set $w_s = f(s,x)$ for $s \le t$, and verify easily that the sequence $(w_0, \dots, w_t)$ satisfies the conditions on the right. For the converse, suppose that $(w_0, \dots, w_t)$ satisfies the conditions on the right and prove by (finite) induction on $s \le t$ that $w_s = f(s,x)$. ⊣

We can view the equivalence (4-3) as a theorem about recursive definitions which have already been justified in some other way; or we can see it as a definition of a function f which satisfies the recursive equations (4-2) and so justifies recursive definitions—which is how Dedekind saw it. He used it to prove that the Peano axioms characterize **N** up to isomorphism, Problem x2.18*. In any case, it reduces proving Theorem 4A.1 to justifying *quantification over finite sequences* within the class of arithmetical relations, and we will do this by an **arithmetical coding** of finite sequences.

The key number-theoretic fact that we need to prove Theorem 4A.1 using Dedekind's analysis is the following classical result. We outline a proof of it using Problems x2.19 and x2.20, which can be assumed as given "black boxes" or derived very easily from the *Fundamental Theorem of Arithmetic*, that *every number greater than* 1 *can be written uniquely* (except for order) *as a product of primes.*

4A.3. **Theorem** (The Chinese Remainder Theorem)**.** *If $d_0, \dots, d_t$ are relatively prime numbers* (i.e., no two of them have a common factor > 1) *and $w_0 < d_0, \dots, w_t < d_t$, then there exists a number a such that*

$$w_0 = \text{rem}(a, d_0), \dots, w_t = \text{rem}(a, d_t).$$

Proof. Consider the set D of all $(t+1)$-tuples bounded by the given numbers $d_0, \dots, d_t$,

$$D = \{(w_0, \dots, w_t) \mid w_0 < d_0, \dots, w_t < d_t\},$$

which has $|D| = d_0 d_1 \cdots d_t$ members, and let

$$A = \{a \mid a < |D|\}$$

which is equinumerous with D. Define the function $\pi : A \to D$ by

$$\pi(a) = (\text{rem}(a, d_0), \text{rem}(a, d_1), \dots, \text{rem}(a, d_t)).$$

Now π is injective (one-to-one), because if $\pi(a) = \pi(b)$ with $a < b < |D|$, then $b - a$ is divisible by each of $d_0, \dots, d_t$ and hence by their product D by Problem x2.20; so $d \le b - a$, which is absurd since $a < b < |D|$. We now apply the **Pigeonhole Principle**: since A and D are equinumerous and $\pi : A \rightarrowtail D$ is an injection, it must be a surjection, and hence whatever $(w_0, \dots, w_t)$ may be, there is an $a < d$ such that

$$\pi(a) = (\text{rem}(a, d_0), \text{rem}(a, d_1), \dots, \text{rem}(a, d_t)) = (w_0, \dots, w_t). \qquad \dashv$$

The idea now is to code an arbitrary tuple $(w_0, \dots, w_t)$ by a pair of numbers (d, a), where d can be used to produce uniformly $t+1$ relatively prime numbers $d_0, \dots, d_t$ and then a comes from the Chinese Remainder Theorem.

4A.4. **Lemma** (Gödel's β-function)**.** *Set*

$$\beta(a, d, i) = \text{rem}(a, 1 + (i+1)d).$$

This is an arithmetical function, and for every tuple of numbers $w_0, \dots, w_t$ there exist numbers a and d such that

$$\beta(a, d, 0) = w_0, \dots, \beta(a, d, t) = w_t.$$

Proof. The β-function is arithmetical because it is defined by substitutions from addition, multiplication and the remainder function, which is arithmetical since (by convention for the case $y = 0$),

$$\text{rem}(x, y) = r \iff \Big[y = 0 \land r = 0\Big] \lor \Big[y \neq 0 \land (\exists q)[x = yq + r \land r < y\Big].$$

To find the required a, d which code the tuple $(w_0, \dots, w_t)$, set

$$s = \max(t + 1, w_0, \dots, w_t), \quad d = s!$$

and verify that the $t + 1$ numbers

$$d_0 = 1 + (0+1)d, d_1 = 1 + (1+1)d, \dots, d_t = 1 + (t+1)d$$

are relatively prime. (If a prime p divides $1+(1+i)s!$ and also $1+(1+j)s!$ with $i < j$, then it must divide their difference $(j-i)s!$, and hence it must divide one of $(j - i)$ or $s!$; in either case, it divides $s!$, since $(j - i) \le s$, and then it must divide 1, since it is assumed to divide $1+(1+i)s!$, which is absurd.) It is also immediate that $w_i < d = s!$, by the definition of s, and so the Chinese Remainder Theorem supplies some a such that

$$(w_0, \dots, w_t) = (\text{rem}(a, d_0), \dots, d_t) = (\beta(a, d, 0), \dots, \beta(a, d, t))$$

as required. $\dashv$

Proof of Theorem 4A.1. By the Dedekind analysis of recursion (4-3) and using the β-function to code tuples, we have

$$f(t,\vec{x}) = w \iff (\exists a)(\exists d)\Big[\beta(a,d,0) = g(\vec{x})$$
$$\wedge (\forall s < t)[\beta(a,d,s+1) = h(\beta(a,d,s),s,\vec{x})] \wedge \beta(a,d,t) = w\Big]$$

Thus the graph of f is arithmetical, by the closure properties of the arithmetical functions and relations in Theorem 3J.1. ⊣

There is no single, standard definition of *rich structure*, but the following notion covers the most important examples:

4B. Structures with tuple coding. A *copy of* $\mathbf{N}$ in a structure $\mathbf{A}$ is a structure $\mathbf{N}' = (\mathbb{N}', 0', S', +', \cdot')$ such that:

1. $\mathbf{N}'$ is isomorphic with the structure of arithmetic $\mathbf{N}$.
2. $\mathbb{N}' \subseteq A$.
3. The set $\mathbb{N}'$, the object $0'$ and the functions $S', +'$ and $\cdot'$ are all $\mathbf{A}$-elementary.

A structure $\mathbf{A}$ *admits tuple coding* if it has a copy of $\mathbf{N}$ and there is a number k and an $\mathbf{A}$-elementary function $\gamma : A^{k+1} \to A$ with the following property: for every tuple $w_0, \dots, w_t \in A$, there is some $\vec{a} \in A^k$ such that

$$\gamma(\vec{a},0) = w_0, \gamma(\vec{a},1) = w_1, \dots, \gamma(\vec{a},t) = w_t,$$

where $0, 1, \dots, t$ are the "$\mathbf{A}$-numbers" $0, 1, \dots, t$ (i.e., the copies of these numbers into A by the given isomorphism of $\mathbf{N}$ with $\mathbf{N}'$).

In this definition, γ plays the role of the β-function in $\mathbf{N}$, and we have allowed for the possibility that triples ($k = 3$), or quadruples ($k = 4$), etc., are needed to code tuples of arbitrary length in A using γ. We might have also allowed the natural numbers to be coded by pairs of elements of A or tweak the definition in various other ways, but this version captures all the examples we will discuss here. The key result is:

4B.1. **Proposition.** *Suppose* $\mathbf{A}$ *admits coding of tuples,*

$$g : A^n \to A, \quad h : A^{n+2} \to A$$

are $\mathbf{A}$*-elementary,* $f : A^{n+1} \to A$*, and for* $t \in N', y \notin N'$*,*

$$f(0,\vec{x}) = g(\vec{x}), \quad f(t+1,\vec{x}) = h(f(t,\vec{x}),t,\vec{x})), \quad f(y,\vec{x}) = y; \tag{4-4}$$

it follows that f *is* $\mathbf{A}$*-elementary.*

This is proved by a simple adjustment of the proof of Theorem 4A.1 and it can be used to show that structures which admit tuple coding have a rich class of elementary functions and relations.

4B.2. **Example** (The integers)**.** The ring of (rational) *integers*

$$\mathbf{Z} = (\mathbb{Z}, 0, 1, +, \cdot) \quad (\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}) \tag{4-5}$$

admits tuple coding.

To see this, we use the fact that $\mathbb{N} \subseteq \mathbb{Z}$, and it is a $\mathbf{Z}$-elementary set because of *Lagrange's Theorem*, by which *every natural number is the sum of four squares*:

$$x \in \mathbb{N} \iff (\exists u, v, s, t)[x = u^2 + v^2 + s^2 + t^2] \quad (x \in \mathbb{Z}).$$

We can then use the β-function (with some tweaking) to code tuples of integers.

4B.3. **Example** (The fractions)**.** The field of *rational numbers* (fractions)

$$\mathbf{Q} = (\mathbb{Q}, 0, 1, +, \cdot)$$

admits tuple coding.

This is a classical theorem of Julia Robinson which depends on a non-trivial, $\mathbf{Q}$-elementary definition of $\mathbb{N}$ within $\mathbb{Q}$. (Robinson's Theorem depends on some results from number theory which are considerably more difficult than Lagrange's Theorem.)

4B.4. **Example** (The real numbers, with $\mathbb{Z}$)**.** The *structure of analysis*

$$(\mathbf{R}, \mathbb{Z}) = (\mathbb{R}, 0, 1, \mathbb{Z}, +, \cdot)$$

admits tuple coding.

This requires some work and it is not a luxury that we have included the integers as a distinguished subset: the *field of real numbers*

$$\mathbf{R} = (\mathbb{R}, 0, 1, +, \cdot)$$

does not admit tuple coding. We will discuss this very interesting, classical structure in the next section.

§5. Quantifier elimination. At the other end of the class of structures which admit tuple coding are some important structures in which the elementary functions and relations are all quite trivial. We will describe some of these and isolate the property of *quantifier eliminability* which makes them "simple"—much as tuple coding makes the structures in the preceding section complex.

5A. Techniques for simplifying formulas. We start with some simple logical equivalences which we will be using, and to simplify notation, we set for arbitrary formulas ϕ, ψ and any structure $\mathbf{A}$:

$$\phi \asymp_{\mathbf{A}} \psi \iff_{\text{df}} \mathbf{A} \models \phi \leftrightarrow \psi,$$
$$\phi \asymp \psi \iff_{\text{df}} \models \phi \leftrightarrow \psi \qquad (\iff \text{ for all } \mathbf{A}, \phi \asymp_{\mathbf{A}} \psi).$$

5A.1. **Proposition** (Basic logical equivalences).

(1) *The distributive laws:*

$$\phi \wedge (\psi \vee \chi) \asymp (\phi \wedge \psi) \vee (\phi \wedge \chi)$$
$$\phi \vee (\psi \wedge \chi) \asymp (\phi \vee \psi) \wedge (\phi \vee \chi)$$

(2) *De Morgan's laws:*

$$\neg(\phi \wedge \psi) \asymp \neg\phi \vee \neg\psi$$
$$\neg(\phi \vee \psi) \asymp \neg\phi \wedge \neg\psi$$

(3) *Double negation, implication and the universal quantifier:*

$$\neg\neg\phi \asymp \phi, \qquad \phi \to \psi \asymp \neg\phi \vee \psi, \quad \forall x\phi \asymp \neg(\exists x)\neg\phi$$

(4) *Renaming of bound variables: if y is a variable which does not occur in ϕ and $\phi\{x :\equiv y\}$ is the result of replacing x by y in all its free occurrences, then*

$$\exists x\phi \asymp \exists y\phi\{x :\equiv y\}$$
$$\forall x\phi \asymp \forall y\phi\{x :\equiv y\}$$

(5) *Distribution law for $\exists$ over $\vee$:*

$$\exists x(\phi_1 \vee \cdots \vee \phi_n) \asymp \exists x\phi_1 \vee \cdots \vee \exists x\phi_n$$

(6) *Pulling the quantifiers to the front: if x does not occur free in ψ, then*

$$\exists x\phi \wedge \psi \asymp \exists x(\phi \wedge \psi)$$
$$\exists x\phi \vee \psi \asymp \exists x(\phi \vee \psi)$$
$$\forall x\phi \wedge \psi \asymp \forall x(\phi \wedge \psi)$$
$$\forall x\phi \vee \psi \asymp \forall x(\phi \vee \psi)$$

PROOF is easy and we leave it for the Problems. ⊣

5A.2. **Corollary** (Prenex normal forms). *Every formula χ is* (effectively) *logically equivalent to a formula*

(5-1) $$\chi^* \equiv Q_1 x_1 \cdots Q_n x_n \psi,$$

where $Q_1, \dots, Q_n$ are quantifiers, ψ is quantifier-free and every variable which occurs free on χ^* also occurs free in χ.

Notation for truth and falsity. For the constructions in this section, it is useful to enrich the language $\mathsf{LPCI}(\tau)$ with propositional constants **tt**, **ff** for truth and falsity. We may think of these as abbreviations,

$$\mathbf{tt} :\equiv \exists x(x = x), \quad \mathbf{ff} :\equiv \forall x(x \neq x),$$

considered (by convention) as prime formulas.

A **literal** is either a prime formula $R(t_1, \dots, t_n)$, $s = t$, **tt**, **ff**, or the negation of a prime formula.

5A.3. **Proposition** (Disjunctive normal form)**.** *Every quantifier-free formula χ is* (effectively) *logically equivalent to a disjunction of conjunctions of literals which has no more variables than χ: i.e., for suitable n, n_i, and literals ℓ_{ij} $(i = 1, \ldots, n, j = 1, \ldots, n_i)$ whose variables all occur in χ,*

$$\chi \asymp \chi^* \equiv \phi_1 \vee \cdots \vee \phi_n, \text{ where for } i = 1, \ldots, n,\ \phi_i \asymp \ell_{i1} \wedge \cdots \wedge \ell_{in_i}.$$

By the definition in the proposition, $x = y \vee \neg(z = z)$ is not a disjunctive normal form of $x = y$ (if all three variables are distinct), even though

$$x = y \asymp x = y \vee \neg(z = z)$$

PROOF. We use Proposition 5A.1 to show that the class $\mathcal{F}$ of formulas for which we can effectively compute a logically equivalent disjunctive normal form contains the literals, because

$$\ell \asymp \ell \vee \mathbf{ff}$$

and is closed under disjunction, conjunction and negation. (The messiest part is the argument that $\mathcal{F}$ is closed under conjunction $\phi \wedge \psi$, which we do by induction on the number of disjuncts in ψ.) ⊣

5B. Proving quantifier eliminability. A *quantifier-free normal form* for a formula χ in a structure $\mathbf{A}$, is any quantifier-free formula χ^* whose variables are among the free variables of χ and such that

$$\chi \asymp_{\mathbf{A}} \chi^*.$$

A structure $\mathbf{A}$ *admits elimination of quantifiers*, if every formula χ has a quantifier-free normal form in $\mathbf{A}$; and it *admits effective elimination of quantifiers*, if there is an effective procedure which will compute for each χ a quantifier-free normal form for χ in $\mathbf{A}$.

Decidable structures. To understand part of the importance of this notion, suppose the vocabulary τ is purely relational, i.e., it has no constant or function symbols. Now the only quantifier-free sentences are **tt** and **ff**; and so if a τ-structure $\mathbf{A}$ admits effective quantifier elimination, then we can effectively decide for each sentence χ whether it is logically equivalent in $\mathbf{A}$ to **tt** or **ff**—in other words, we have a *decision procedure for truth in* $\mathbf{A}$.

More generally, suppose τ may have constants and function symbols and $\mathbf{A}$ admits effective quantifier elimination: if we have a decision procedure for quantifier-free *sentences* (with no variables), then we have a decision procedure for truth in $\mathbf{A}$. The hypothesis is, in fact, satisfied in most examples of structures which admit quantifier elimination.

5B.1. **Lemma** (**Quantifier elimination test**)**.** *If every formula of the form*

$$\chi \equiv \exists x[\chi_1 \wedge \cdots \wedge \chi_n] \quad (\textit{where } \chi_1, \ldots, \chi_n \textit{ are literals})$$

is effectively equivalent in a structure **A** *to a quantifier-free formula whose variables are all among the free variables of* χ*, then* **A** *admits effective quantifier elimination.*

Proof. It is enough by (3) of Proposition 5A.1 to eliminate quantifiers for formulas which don't have any implications or universal quantifiers. Thus, if $\mathcal{F}$ is the class of all such formulas which (effectively) have quantifier free forms in **A**, then it is enough to show that it contains all literals, which it clearly does; that it is closed under $\neg, \wedge$ and $\vee$, which it clearly is; and that it is closed under existential quantification. For the latter, if

$$\chi \equiv \exists x \phi$$

with ϕ quantifier-free, we bring ϕ to disjunctive normal form, so that

$$\chi \asymp \exists x[\phi_1 \vee \cdots \vee \phi_n] \asymp \exists x\phi_1 \vee \cdots \vee \exists x\phi_n$$

where each ϕ_i is a conjunction of literals and then we use the hypothesis of the Lemma. ⊣

5B.2. **Proposition.** *For each infinite set* A*, the structure* $\mathbf{A} = (A)$ *in the language with empty vocabulary admits effective quantifier elimination.*

Proof. By the test 5B.1, it is enough to eliminate quantifiers from every formula of the form

$$\chi \asymp \exists x\Big[(x = z_1 \wedge \cdots \wedge x = z_k) \wedge (u_1 = v_1 \wedge \cdots u_l = v_l) \\ \wedge (x \neq w_1 \wedge \cdots \wedge x \neq w_m) \wedge (s_1 \neq t_1 \wedge \cdots \wedge t_o \neq s_o)\Big]$$

where we have grouped the variable equations and inequations according to whether x occurs in them or not; i.e., x is none of the variables u_i, v_i, s_i, t_i. We can also assume that x is none of the variables z_i, because the equation $x = x$ can simply be deleted; and it is none of the variables w_i, since if $x \neq x$ is one of the conjuncts, then $\chi \asymp \mathbf{ff}$.

Case 1, $k = 0$, i.e., there is no equation of the form $x = z$ in the matrix of χ. In this case

$$\chi \asymp (u_1 = v_1 \wedge \cdots u_l = v_l) \wedge (s_1 \neq t_1 \wedge \cdots \wedge t_m \neq s_m).$$

This is because if π is any assignment which satisfies

$$(u_1 = v_1 \wedge \cdots u_l = v_l) \wedge (s_1 \neq t_1 \wedge \cdots \wedge t_m \neq s_m)$$

and t is any element in the (infinite) set A which is distinct from $\pi(w_1), \ldots$, $\pi(w_m)$, then $\pi\{x := t\}$ satisfies the matrix of χ.

Case 2, $k > 0$, so there is an equation $x = z_i$ in the matrix of χ. In this case,

$$\chi \asymp (x = z_1 \wedge \cdots \wedge x = z_k)\{x :\equiv z_i\} \wedge (u_1 = v_1 \wedge \cdots u_l = v_l) \\ \wedge (x \neq w_1 \wedge \cdots \wedge x \neq w_m)\{x :\equiv z_i\} \wedge (s_1 \neq t_1 \wedge \cdots \wedge t_m \neq s_m)]$$

since every assignment which satisfies χ must assign to x the same value that it assigns to z_i. ⊣

This proposition is about a structure of no interest whatsoever, but the method of proof is typical of many quantifier elimination proofs.

5B.3. **Definition** (Dense linear orderings)**.** A linear ordering $\mathbf{L} = (L, \leq)$ is *dense in itself* if for every $x, y \in L$ such that $x < y$, there is a z such that $x < z < y$.

Standard examples are the usual orderings $(\mathbb{Q}, \leq)$ and $(\mathbb{R}, \leq)$ on the rational and the real numbers. They also have no least or greatest element, so they are covered by the next result.

5B.4. **Theorem.** *If* $\mathbf{L} = (L, \leq)$ *is a dense linear ordering without least or greatest element, then* $\mathbf{L}$ *admits effective quantifier elimination.*

Outline or proof. It is convenient to introduce a new symbol $<$ for strict inequality, so that

$$\text{(5-2)} \quad \mathbf{ff} \asymp_{\mathbf{L}} x < x, \quad \mathbf{tt} \asymp_{\mathbf{L}} x = x, \quad x \leq y \asymp_{\mathbf{L}} x = y \vee x < y, \\ x < y \asymp_{\mathbf{L}} x \leq y \wedge x \neq y.$$

We can use the equivalences in the first line to eliminate the truth values and the symbol $\leq$, so that every formula is logically equivalent in $\mathbf{L}$ to one in which only the symbols $=$ and $<$ occur. In particular, the literals which occur in disjunctive normal forms of quantifier free formulas are all in one of the forms

$$x = y, \quad x \neq y, \quad x < y, \quad \neg(x < y)$$

We now replace all the negated literals by quantifier free formulas which have no negation, using the equivalences

$$\text{(5-3)} \quad x \neq y \asymp_{\mathbf{L}} x < y \vee y < x, \quad \neg(x \leq y) \asymp_{\mathbf{L}} y < x, \\ \neg(x < y) \asymp_{\mathbf{L}} x = y \vee y < x$$

and then we apply repeatedly the Distributive Laws in Proposition 5A.1 (which do not introduce negations) to construct a disjunctive normal form with only positive literals $x = y$ and $x < y$. This means that in applying the basic test Lemma 5B.1, we need consider only formulas of the form

$$\begin{aligned}\chi \equiv \exists x\Big[&(x = z_1 \wedge \cdots \wedge x = z_k)\\ &\wedge (x < u_1 \wedge \cdots \wedge x < u_l) \wedge (v_1 < x \wedge \cdots \wedge v_m < x)\\ &\wedge (s_1 < s'_1 \wedge \cdots \wedge s_n < s'_n) \wedge (t_1 = t'_1 \wedge \cdots \wedge t_o = t'_o)\Big]\end{aligned}$$

If some $u_i \equiv x$ or some $v_j \equiv x$, then $\chi \asymp_{\mathbf{L}} \mathbf{ff}$, so we may assume that these variables are all distinct from x.

Case 1, $k > 0$, so that some equation $x = z_i$ is present in the matrix. Now χ is equivalent to the quantifier-free formula which is constructed by replacing x by z_i in the matrix.

Case 2, $k = l = m = 0$, so that x does not occur in the matrix of χ. We simply delete the quantifier.

Case 3, $k = l = 0$ but $m > 0$. In this case

$$\chi \asymp_{\mathbf{L}} (s_1 < s'_1 \wedge \cdots \wedge s_n < s'_n) \wedge (t_1 = t'_1 \wedge \cdots \wedge t_o = t'_o)]$$

because whatever values are assigned to $v_1, \ldots, v_m$ by an assignment, some greater value can be assigned to x since $\mathbf{L}$ has no largest element.

Case 4, $k = m = 0$ but $l > 0$. This case is symmetric to Case 3, and we handle it using the fact that $\mathbf{L}$ has no least element.

Case 5, $k = 0$ but $m > 0, l > 0$. Since $\mathbf{L}$ is dense in itself, the restrictions on x in the matrix will be satisfied by some x exactly when

$$\max\{v_1, \ldots, v_m\} < \min\{u_1, \ldots, u_l\},$$

and we can say this formally by a big conjunction: i.e.,

$$\begin{aligned}\chi \asymp_{\mathbf{L}} &\textstyle\bigwedge\!\!\!\bigwedge_{1 \le i \le l, 1 \le j \le m}(v_j < u_i)\\ &\wedge (s_1 < s'_1 \wedge \cdots \wedge s_n < s'_n) \wedge (t_1 = t'_1 \wedge \cdots \wedge t_o = t'_o)\end{aligned}$$

This completes the verification of the test, Lemma 5B.1 for dense linear orderings with no first and last element, and so these structures admit effective quantifier elimination. ⊣

5C. Additional examples. We list some of the most interesting examples of structures which admit effective quantifier elimination without proofs—which can be found in many standard textbooks in logic.

5C.1. **Example** (Successor arithmetic)**.** The reduct $(\mathbb{N}, 0, S)$ of $\mathbf{N}$ without addition or multiplication admits effective quantifier elimination, as does the somewhat richer structure $(\mathbb{N}, 0, S, <)$.

5C.2. **Example** (Presburger arithmetic)**.** The reduct $(\mathbb{N}, 0, S, +)$ of $\mathbf{N}$ does not quite admit quantifier elimination, but something quite close to it does. For each $m \ge 2$, let

$$x \equiv_m y \iff m \text{ divides } y - x \quad (x \text{ is congruent to } y \text{ mod } m),$$

and consider (for this one time) the expansion of $(\mathbb{N}, 0, S, +)$ by these infinitely many relations,

$$\mathbf{N}_P = (\mathbb{N}, 0, S, +, \{\equiv_m\}_{m \in \mathbb{N}, m \geq 2}).$$

This structure admits effective quantifier elimination and there is a trivial decision procedure for quantifier free sentences, which involve only numerals and congruence assertions about them; and so it is a decidable structure, and then the structure $(\mathbb{N}, 0, S, +)$ of additive arithmetic is also decidable, since it is a reduct of $\mathbf{N}_P$.

This is a famous, non-trivial theorem of Presburger.

Note that the expansion of the language by these congruence relations is quite similar to the expansion with **tt** and **ff** which we have already assumed, because we need it. The congruence relations are simply definable in additive arithmetic, one-at-a-time:

$$x \equiv_m y :\equiv (\exists z)\Big[(x + \underbrace{z + z + \cdots + z}_{m \text{ times}} = y) \vee (y + \underbrace{z + z + \cdots + z}_{m \text{ times}} = x)\Big].$$

The quantifier elimination in Presburger's structure $\mathbf{N}_P$ yields for each χ a quantifier-free formula in which these new, prime formulas $x \equiv_m y$ occur, for various values of m; we can then replace all of them with their definition, which gives us a formula χ^* which is $\asymp_{\mathbf{N}_P}$ with χ and in which existential quantifiers occur only in the "literals". This is exactly the sort of "extended quantifier-free" formulas that we will get if we replace **tt** and **ff** by their definitions after the quantifier elimination procedure has been completed.

5C.3. **Example** (The complex numbers)**.** The field of complex numbers

$$\mathbf{C} = (\mathbb{C}, 0, 1, +, \cdot)$$

admits effective quantifier elimination, and so it is decidable, since the quantifier-free sentences in the language involve only trivial equalities and inequalities about numerals.

5C.4. **Example** (The ordered field of real numbers)**.** The structure

$$\mathbf{R}_o = (\mathbb{R}, 0, 1, +, \cdot, \leq)$$

admits effective quantifier elimination, and so it is decidable, as above.

This is a famous, deep theorem of Tarski, especially important because it establishes the decidability of classical (ancient) Euclidean plane and space geometry: it is easy to see that if we use Cartesian coordinates, we can translate all the propositions studied in Euclidean geometry into sentences in the language of $\mathbf{R}_o$, and then decide them by Tarski's algorithm. Contrast this result with Example 4B.4: if we just add a name for the set of integers $\mathbb{Z}$ to the language, we get a structure which admits tuple coding, in whose language we can express all the propositions of classical analysis—including calculus.

§6. Theories and elementary classes. Up until now we have been studying the interpretation of single sentences of $\mathsf{LPCI}(\tau)$ on a single τ-structure; next we consider classes of structures and investigate how sets of sentences can be used to define their elementary properties.

6A. Logical consequence. A **theory** in a language $\mathsf{LPCI}(\tau)$ is any (possibly infinite) set of sentences T of $\mathsf{LPCI}(\tau)$. The members of T are its *axioms*.

A τ-structure $\mathbf{A}$ is a *model* of T if every sentence of T is true in $\mathbf{A}$: we write

$$\mathbf{A} \models T \iff_{\text{df}} \text{for all } \phi \in T,\ \mathbf{A} \models \phi, \tag{6-1}$$

and we collect all the models of T into a class of structures,

$$\text{Mod}(T) =_{\text{df}} \{\mathbf{A} \mid \mathbf{A} \models T\}. \tag{6-2}$$

A property Φ of τ-structures is *elementary* if there is a τ-theory T such that

$$\mathbf{A} \text{ has property } \Phi \iff \mathbf{A} \models T, \tag{6-3}$$

and Φ is *basic elementary* if (6-3) holds with a finite T—or, equivalently if for some sentence χ,

$$\mathbf{A} \text{ has property } \Phi \iff \mathbf{A} \models \chi \qquad (\text{with } \chi \equiv \textstyle\bigwedge\!\!\!\bigwedge T). \tag{6-4}$$

We will also "identify" a property of structures with its extension and write synonymously

$$\mathbf{A} \text{ has property } \Phi \iff \mathbf{A} \in \Phi;$$

and then *elementary* and *basic elementary* classes of structures are characterized by the conditions

$$\begin{aligned} &\mathbf{A} \in \Phi \iff \text{for every } \phi \in T, \mathbf{A} \models \phi && (\textbf{elementary class}) \\ &\mathbf{A} \in \Phi \iff \mathbf{A} \models \chi && (\textbf{basic elementary class}) \end{aligned}$$

In the opposite direction, the *theory of a τ-structure* $\mathbf{A}$ is the set of all $\mathsf{LPCI}(\tau)$-sentences that it satisfies,

$$\text{Th}(\mathbf{A}) =_{\text{df}} \{\chi \mid \chi \text{ is a sentence and } \mathbf{A} \models \chi\}; \tag{6-5}$$

and two τ-structures are **elementarily equivalent** if they satisfy the same sentences,

$$\mathbf{A} \approx \mathbf{B} \iff_{\text{df}} \text{Th}(\mathbf{A}) = \text{Th}(\mathbf{B}). \tag{6-6}$$

It is easy to check that *isomorphic structures are elementarily equivalent*

$$\mathbf{A} \cong \mathbf{B} \Longrightarrow \mathbf{A} \approx \mathbf{B},$$

cf. Problem x2.33*. We will see that the converse fails quite spectacularly, even when $\mathbf{A}$ is the structure of arithmetic $\mathbf{N}$, cf. Section 7E.

Finally, we define the fundamental relation of **logical consequence** between a theory and a sentence,

(6-7) $$T \models \chi \iff \text{for every } \mathbf{A},\ \text{if } \mathbf{A} \models T,\ \text{then } \mathbf{A} \models \chi;$$

and a τ-sentence χ is **valid** if it is a logical consequence of the empty theory,

(6-8) $$\chi \text{ is valid } \iff \models \chi \iff \text{ for every } \tau\text{-structure } \mathbf{A}, \mathbf{A} \models \chi.$$

6B. Some basic theories. We collect here for easy reference the definitions of a few important theories.

6B.1. **Symmetric graphs**. The theory SG of *symmetric graphs* is formulated in the language $\mathsf{LPCI}(E)$ with just one, binary relation symbol E and two axioms,

$$\forall x \neg E(x,x), \quad \forall x \forall y [E(x,y) \leftrightarrow E(y,x)].$$

The symmetric graphs then are exactly the models of SG.

6B.2. **Partial orders**. The theories of *partial* and *linear orderings* are also formulated in the language with vocabulary just one, binary relation symbol, this time $\leq$:

$$\mathsf{PO} =_{\text{df}} \Big\{\forall x(x \leq x),\ \forall x \forall y[(x \leq y \wedge y \leq x) \rightarrow x = y], \\ \forall x \forall y \forall z[(x \leq y \wedge y \leq z) \rightarrow x \leq z]\Big\}$$

$$\mathsf{LO} =_{\text{df}} \mathsf{PO} \cup \Big\{\forall x \forall y[x \leq y \vee y \leq x]\Big\}$$

$$\mathsf{DLO} =_{\text{df}} \mathsf{LO} \cup \Big\{\forall x \forall y[x < y \rightarrow \exists z(x < z \wedge z < y)], \\ \forall x \exists y[x < y],\ \forall y \exists x[x < y]\Big\}$$

This last DLO is the theory of linear orderings which are dense in themselves and have no least or greatest element, e.g., $(\mathbb{Q}, \leq)$, cf. Theorem 5B.4.

6B.3. **Fields**. The theory Fields comprises the formal expressions of the axioms for a field listed in Definition 1C, in the language $\mathsf{LPCI}(0,1,+,\cdot)$.

For each number $n \geq 1$ define the term $n \cdot 1$ by the recursion

$$1 \cdot 1 :\equiv 1, \quad (n+1) \cdot 1 :\equiv (n \cdot 1) + (1),$$

so that e.g., $3 \cdot 1 \equiv ((1) + (1)) + (1)$. (Make sure you understand here what is a term of $\mathsf{LPCI}(0,1,+,\cdot)$, what is an ordinary number, and which $+$ is meant in the various places.)

For each prime number p, the finite set of sentences

$$\mathsf{Fields}_p =_{\text{df}} \mathsf{Fields} \cup \Big\{\neg(2 \cdot 1 = 0),\ \dots, \neg((p-1) \cdot 1 = 0),\ p \cdot 1 = 0\Big\}$$

is the theory of *fields of characteristic* p. The theory of *fields of characteristic* 0 is defined by

$$\mathsf{Fields}_0 =_{\mathrm{df}} \mathsf{Fields} \cup \Big\{ \neg(2\cdot 1 = 0),\ \neg(3\cdot 1 = 0), \dots \Big\}.$$

The simplest example of a field of characteristic p is the finite structure

$$\mathbf{F}_p = (\{0, 1, \dots, p-1\}, 0, 1, +, \cdot)$$

with the usual operations on it executed *modulo* p, but it takes some (algebra) work to show that this is a field. There are many other fields of characteristic p, both finite and infinite. The number fields $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{C}$ have characteristic 0.

For each prime $p \in \mathbb{N}$, Fields_p is a finite theory while Fields_0 is infinite.

To simplify the definitions of theories we often use the *universal closures* of formulas, i.e., the sentences

$$\vec{\forall}\,\phi \equiv_{\mathrm{df}} \forall \mathrm{v}_0 \forall \mathrm{v}_1 \dots \forall \mathrm{v}_n \phi$$

where n is least so that all the free variables of ϕ are among $\mathrm{v}_0, \dots, \mathrm{v}_n$.

6B.4. **Peano Arithmetic, PA.** The axioms of Peano arithmetic are the natural LPCI-formalizations of the Peano axioms for $\mathbf{N}$ listed in section 1D, i.e., the universal closures of the following formulas in the language $\mathsf{LPCI}(0, S, +, \cdot)$:

1. $\neg[S(x) = 0]$.
2. $S(x) = S(y) \rightarrow x = y$.
3. $x + 0 = x,\ x + S(y) = S(x + y)$.
4. $x \cdot 0 = 0,\ x \cdot (Sy) = x \cdot y + x$.
5. For every formula $\phi(x, \vec{y})$,

$$\Big[\phi(0, \vec{y}) \wedge (\forall x)[\phi(x, \vec{y}) \rightarrow \phi(S(x), \vec{y})]\Big] \rightarrow (\forall x)\phi(x, \vec{y}).$$

The last is the **Elementary Axiom Scheme of Induction** which approximates in LPCI the intended meaning of the full Axiom of Induction in Definition 1D. It has infinitely many *instances*, one for each formula $\phi(x, \vec{y})$, and so PA is an infinite theory.

6B.5. **The (Raphael) Robinson system Q**. This is a weak, finite theory of natural numbers in the language of arithmetic, which replaces the Induction Scheme by the single claim that every non-zero number is a successor. Its axioms are the universal closures of the following formulas:

1. $\neg[S(x) = 0]$.
2. $S(x) = S(y) \rightarrow x = y$.
3. $x + 0 = x,\ x + (S(y) = S(x + y)$.
4. $x \cdot 0 = 0,\ x \cdot (Sy) = x \cdot y + x$.
5. $x = 0 \vee (\exists y)[x = S(y)]$.

It is clear that **N** is a model of Q, but Q is a weak theory, and so it is quite easy to construct many peculiar models of it.

6C. Quantifier elimination for theories. A theory T in the language LPCI(τ) *admits elimination of quantifiers*, if for every τ-formula χ, there is a quantifier-free formula χ^* (whose variables are all among the free variables of χ) such that

$$T \models \chi \leftrightarrow \chi^*.$$

As with structures, we assume here that the language is expanded by the prime, propositional constants **tt** and **ff** which may occur in χ^*.

Notice that a structure **A** admits (effective) quantifier elimination exactly when its theory Th(**A**) admits (effective) quantifier elimination.

6C.1. **Theorem.** *The theory* DLO *of dense linear orderings without first or last element admits effective quantifier elimination, and so it is decidable.*

PROOF follows immediately from the proof of Theorem 5B.4, which produces the same quantifier-free form **L**-equivalent to a given χ, independently of the specific **L**, just so long as **L** $\models$ DLO.

The second claim simply means that we can decide for any given sentence χ whether or not DLO $\models \chi$. It is true because the quantifier elimination procedure yields either **tt** or **ff** as **L**-equivalent to χ, independently of the specific **L**. $\dashv$

§7. Formal deduction. The (Hilbert style) proof system for LPCI is the natural extension to LPCI of the system of axioms and rules for the propositional calculus. In the notation we will use here:

- t stands for an arbitrary term of LPCI(τ), for a fixed τ;
- ϕ, ψ, χ stand for arbitrary formulas of LPCI(τ); and
- $\phi(v, \vec{u})$ stands for an extended formula, so that all the free variables of ϕ are in the list $v, \vec{u}$.

Some of the axioms and rules have a condition, the most significant being the hypothesis that

$$t \textit{ is free for } v \textit{ in } \phi(v, \vec{u});$$

it means that

no occurrence of a variable in t is bound in the formula $\phi(t, \vec{u})$.

7A. The Hilbert axiomatic system. The axiom schemes and rules for LPCI(τ) are listed in Diagram 1. They are naturally divided into four groups, the first three of which ((**A**) – (**C**)) specify LPC(τ), the lower predicate calculus without identity.

(A) Propositional axiom schemes, same as in PL.

(1) $\phi \to (\psi \to \phi)$
(2) $(\phi \to \psi) \to ((\phi \to (\psi \to \chi)) \to (\phi \to \chi))$
(3) $(\phi \to \psi) \to ((\phi \to \neg\psi) \to \neg\phi)$
(4) $\neg\neg\phi \to \phi$
(5) $\phi \to (\psi \to (\phi \wedge \psi))$
(6a) $(\phi \wedge \psi) \to \phi$ (6b) $(\phi \wedge \psi) \to \psi$
(7a) $\phi \to (\phi \vee \psi)$ (7b) $\psi \to (\phi \vee \psi)$
(8) $(\phi \to \chi) \to ((\psi \to \chi) \to ((\phi \vee \psi) \to \chi))$

(B) Predicate axiom schemes.

(9) $\forall v\phi(v, \vec{u}) \to \phi(t, \vec{u})$ (t free for v in $\phi(v, \vec{u})$)
(10) $\forall v(\phi \to \psi) \to (\phi \to \forall v\psi)$ (v not free in ϕ)
(11) $\phi(t, \vec{u}) \to \exists v\phi(v, \vec{u})$ (t free for v in $\phi(v, \vec{u})$)

(C) Rules of inference.

(12) From ϕ and $\phi \to \psi$, infer ψ. (Modus Ponens)
(13) From ϕ, infer $\forall v\phi$. (Generalization)
(14) From $\phi \to \psi$, infer $\exists v\phi \to \psi$, provided v is not free in ψ. (Exists Elimination)

(D) Identity axioms.

(15) $v = v \quad v = v' \to v' = v \quad v = v' \to (v' = v'' \to (v = v''))$
(16) $(v_1 = w_1 \wedge \dots v_n = w_n) \to (R(v_1, \dots, v_n) \to R(w_1, \dots, w_n))$ (R n-ary relation symbol)
(17) $(v_1 = w_1 \wedge \dots v_n = w_n) \to (f(v_1, \dots, v_n) = f(w_1, \dots, w_n))$ (f n-ary function symbol)

DIAGRAM 1. The Hilbert system for LPCI

A **deduction** in LPCI from a theory T is any sequence of formulas

$$\phi_0, \phi_1, \dots, \phi_n,$$

where each ϕ_i is either an axiom, or a sentence in T, or the same as some ϕ_j for some $j < i$, or follows from previously listed formulas by one of the rules of inference. We set

$$T \vdash \phi \iff_{\text{df}} \text{there exists a deduction } \phi_0, \dots, \phi_n \text{ from } T \text{ with } \phi \equiv \phi_n.$$

If $T = \emptyset$ we just write $\vdash \phi$. A **proof** is a deduction from the empty theory.

A deduction is **propositional** if it only uses the propositional axioms in group (A) and the Modus Ponens inference rule (12). (The formulas in a propositional deduction may have the identity symbol "=" and quantifiers in them.)

If T is a theory, χ is a sentence and $T \vdash \chi$, we call χ a proof theoretic consequence or just a (formal) **theorem** of T. A **propositional theorem** of T is any formula ϕ for which there is a propositional deduction from T; these are easy to recognize, because they are substitution instances of propositional tautologies, cf. Problem x2.41.

The proof theoretic consequences of the empty theory are the theorems of $\mathsf{LPCI}(\tau)$.

7B. The Soundness Theorem for LPCI. We now turn to the first of the two basic facts which relate the syntax and the semantics of LPCI:

7B.1. **Theorem.** *Every theorem of a theory T is a logical consequence of T, i.e., for every sentence χ,*

(7-1) $$\text{if } T \vdash \chi, \text{ then } T \models \chi.$$

This is a technical result whose proof can be omitted in a first introduction to logic, but it is worth including here for two reasons: it illustrates some of the methods of proof used in **Proof Theory**, one of the basic parts of our subject, and it explains the meaning of—and the reasons for including—the restrictions in the predicate axiom schemes and inference rules of LPCI.

7B.2. **Lemma** (Soundness of the axioms)**.** *If χ is any instance of the axiom schemes* (1) – (11) *and* (15) – (17) *of* $\mathsf{LPCI}(\tau)$ *and* $\mathbf{A}$ *is any τ-structure, then* $\mathbf{A}$ *satisfies the universal closure of χ,* $\mathbf{A} \models \vec{\forall}\chi$.

Proof. First we verify easily, directly from the definitions and the Compositionality Theorem 3G.1 that

(7-2) $$\mathbf{A} \models \vec{\forall}\chi \iff \text{for every assignment } \pi, \mathbf{A}, \pi \models \chi;$$

and then we prove by direct computation that if χ is an instance of one of the axiom schemes (1) – (11), (15) – (17), then $\mathbf{A}, \pi \models \chi$, for every assignment π. The argument is trivial for the propositional and identity axioms (1) – (8) and (15) – (17). For the somewhat more complex (9) – (11) it is convenient to prove first two sublemmas.

Sublemma 1. For every term α, variable v, term t and assignment π,

(∗) $$\text{value}^{\mathbf{A}}(\alpha\{v :\equiv t\}, \pi) = \text{value}^{\mathbf{A}}(\alpha, \pi\{v := \text{value}^{\mathbf{A}}(t, \pi)\}).$$

Proof. We fix $\mathbf{A}$, t and π and prove (∗) by induction on the term α.
If v does not occur in α then the two sides of (∗) are identical; and in the basis case, if $\alpha \equiv v$, then

$$\text{value}^{\mathbf{A}}(v\{v :\equiv t\}, \pi) = \text{value}^{\mathbf{A}}(t, \pi) = \text{value}^{\mathbf{A}}(v, \pi\{v := \text{value}^{\mathbf{A}}(t, \pi)\}.$$

For the inductive step we compute, using the definition of $\mathrm{value}^{\mathbf{A}}(s, \pi)$:

$$\begin{aligned}
&\mathrm{value}^{\mathbf{A}}(f(s_1, \dots, s_n)\{v :\equiv t\}, \pi) \\
&= \mathrm{value}^{\mathbf{A}}(f(s_1\{v :\equiv t\}, \dots, s_n\{v :\equiv t\}), \pi) \\
&= f^{\mathbf{A}}(\mathrm{value}^{\mathbf{A}}(s_1\{v :\equiv t\}, \pi), \dots, \mathrm{value}^{\mathbf{A}}(s_n\{v :\equiv t\}, \pi)) \\
&= f^{\mathbf{A}}(\mathrm{value}^{\mathbf{A}}(s_1, \pi\{v : \mathrm{value}^{\mathbf{A}}(t, \pi)\}), \dots, \mathrm{value}^{\mathbf{A}}(s_n, \pi\{v : \mathrm{value}^{\mathbf{A}}(t, \pi)\}) \\
&= \mathrm{value}^{\mathbf{A}}(f(s_1, \dots, s_n), \pi\{v : \mathrm{value}^{\mathbf{A}}(t, \pi)\}),
\end{aligned}$$

where the induction hypothesis is used to infer the fourth from the third line. ⊣ (Sublemma 1)

Sublemma 2. For every formula ϕ, variable v, term t which is free for v in ϕ and assignment π,

$$(**) \qquad \mathbf{A}, \pi \models \phi\{v :\equiv t\} \iff \mathbf{A}, \pi\{v := \mathrm{value}^{\mathbf{A}}(t, \pi)\} \models \phi.$$

Proof. We fix again $\mathbf{A}$ and t and we prove that $(**)$ *holds for every assignment* π by induction on the formulas ϕ .

The argument is very easy for prime formulas, e.g.,

$$\begin{aligned}
&\mathbf{A}, \pi \models (s_1 = s_2)\{v :\equiv t\} \\
&\iff \mathrm{value}^{\mathbf{A}}(s_1\{v :\equiv t\}, \pi) = \mathrm{value}^{\mathbf{A}}(s_2, \{v :\equiv t\}, \pi) \\
&\iff \mathrm{value}^{\mathbf{A}}(s_1, \pi\{v := \mathrm{value}^{\mathbf{A}}(t, \pi)\}) = \mathrm{value}^{\mathbf{A}}(s_2, \pi\{v := \mathrm{value}^{\mathbf{A}}(t, \pi)\}) \\
&\iff \mathbf{A}, \pi\{v := \mathrm{value}^{\mathbf{A}}(t, \pi)\}) \models s_1 = s_2,
\end{aligned}$$

where Sublemma 1 is used to infer the third from the second line.

The argument is also routine in the induction step when ϕ is a propositional combination of smaller formulas.

Suppose then that $\phi \equiv \exists u\psi$; now $(**)$ is trivial if v does not occur in ψ, so we assume it does, and then the hypothesis that t is free for v in ϕ guarantees that u *does not occur in* t. With these assumptions then and the induction hypothesis, we compute:

$$\begin{aligned}
&\mathbf{A}, \pi \models \exists u\psi\{v \equiv t\} \iff \text{for some } x, \mathbf{A}, \pi\{u := x\} \models \psi\{v :\equiv t\} \\
&\iff \text{for some } x, \mathbf{A}, \pi\{u := x\}\{v := \mathrm{value}^{\mathbf{A}}(t, \pi\{u := x\} \models \psi \\
&\qquad \text{(by the induction hypothesis on the assignment } \pi\{u := x\}) \\
&\iff \text{for some } x, \mathbf{A}, \pi\{u := x\}\{v := \mathrm{value}^{\mathbf{A}}(t, \pi) \models \psi \\
&\text{(because } \mathrm{value}^{\mathbf{A}}(t, \pi\{u := x\}) = \mathrm{value}^{\mathbf{A}}(t, \pi) \text{ since } u \text{ does not occur in } t) \\
&\iff \text{for some } x, \mathbf{A}, \pi\{v := \mathrm{value}^{\mathbf{A}}(t, \pi)\{u := x\} \models \psi \\
&\iff \mathbf{A}, \pi\{v := \mathrm{value}^{\mathbf{A}}(t, \pi) \models \exists u\psi.
\end{aligned}$$

The argument for $\phi \equiv \forall u\psi$ is similar. ⊣ (Sublemma 2)

We now return to the proof of the Lemma for Axiom Schemes (9) – (11)

Axiom (9): If t is free for v in ϕ, then $\mathbf{A}, \pi \models \forall v\phi \to \phi\{v :\equiv t\}$. It is enough to prove that

$$\text{if } \mathbf{A}, \pi \models \forall v\phi, \text{ then } \pi \models \phi\{v :\equiv t\},$$

and for this we compute:

$$\begin{aligned} \mathbf{A}, \pi \models \forall v\phi &\Longrightarrow \text{for every } x, \mathbf{A}, \pi\{v := x\}\phi \\ &\Longrightarrow \mathbf{A}, \pi\{v := \text{value}^{\mathbf{A}}(t, \pi)\} \models \phi \Longrightarrow \mathbf{A}, \pi \models \phi\{v :\equiv t\}, \end{aligned}$$

the last inference by Sublemma 2.

Axiom (10): If v does not occur free in ϕ, then

$$\mathbf{A}, \pi \models \forall v(\phi \to \psi) \to (\phi \to \forall v\psi)$$

It is enough to show that if an assignment π satisfies $\forall v(\phi \to \psi)$ (in $\mathbf{A}$), then it also satisfies $(\phi \to \forall v\psi)$; i.e., if π satisfies $\forall v(\phi \to \psi)$ and ϕ, then it satisfies $\forall v\psi$. The hypothesis means that for every $x \in A$,

$$\mathbf{A}, \pi\{v := x\} \models \phi \to \psi;$$

the restriction that v does not occur in ϕ means that

$$\mathbf{A}, \pi \models \phi \iff \mathbf{A}, \pi\{v := x\} \models \phi,$$

by (2) of the Compositionality Theorem 3G.1; and so the assumption that $\mathbf{A}, \pi \models \phi$ implies that

$$\mathbf{A}, \pi\{v := x\} \models \psi.$$

So we have shown that for every $x \in A$, $\mathbf{A}, \pi\{v := x\} \models \psi$, which means exactly that $\mathbf{A}, \pi \models \forall v\psi$.

We skip the argument for Axiom (11) which is very similar to that for Axiom (9). ⊣

More interesting than these straight verifications of the axiom schemes are the counterexamples which show that the restrictions in (9) – (11) are necessary.

7B.3. Axiom scheme (9)

$$\forall v\phi(v, \vec{u}) \to \phi(t, \vec{u}) \qquad (t \text{ free for } v \text{ in } \phi(v, \vec{u}))$$

is not valid in every structure without the restriction. A counterexample is the instance in the language of arithmetic

$$\forall v \exists y[y = S(v)] \to \exists y[y = S(t)]$$

which with $t :\equiv y$ becomes

$$\forall v \exists y[y = S(v)] \to \exists y[y = S(y)],$$

which is clearly false in $\mathbf{N}$.

Axiom scheme (11)

$$\phi(t, \vec{u}) \to \exists v \phi(v, \vec{u}) \qquad (t \text{ free for } v \text{ in } \phi(v, \vec{u}))$$

is not valid without the restriction. A counterexample which uses just the identity symbol is $\phi(v, \vec{u}) \equiv \forall s[s = v]$. Now the instance of (11) is

$$\forall s[s = t]) \to \exists v \forall s[s = v];$$

and if we substitute (the forbidden) $t :\equiv s$, we get

$$\forall s[s = s] \to \exists v \forall s[s = v],$$

which is clearly false in every structure which has more than one element.

We leave (10) for Problem x2.38.

Proof of the Soundness Theorem 7B.1. It is enough to show that if $\phi_0, \dots, \phi_n$ is a deduction from T, then for every structure $\mathbf{A} \models T$ and every assignment π,

$$\mathbf{A}, \pi \models \phi_i \qquad (i = 0, \dots, n)$$

and we do this by (complete) induction on $i \le n$. The conclusion is true when ϕ_i is an axiom by Lemma 7B.2 and it is also trivial if ϕ occurs earlier in the deduction or is inferred from formulas earlier in the deduction by Modus Ponens. Thus it is enough to consider only the case when ϕ is inferred from a formula preceding it in the proof by one of the predicate rules of inference (13) or (14).

(13), Generalization. By the induction hypothesis, for every assignment ρ, we know that $\mathbf{A}, \rho \models \phi$, and we must show that for every assignment π, $\mathbf{A}, \pi \models \forall v \phi$, i.e.,

$$\text{for every } x \in A,\ \mathbf{A}, \pi\{v := x\} \models \phi;$$

but this follows immediately by applying the induction hypothesis to $\rho = \pi\{v := x\}$.

(14) Exists Elimination. By the induction hypothesis, for every assignment ρ, we know that

$$\mathbf{A}, \rho \models \phi \to \psi,$$

and we must show that for every π,

$$\mathbf{A}, \pi \models \exists v \phi \to \psi.$$

If $\mathbf{A}, \pi \not\models \exists v \phi$, then we are done, by the definition of satisfaction for material implication.

If $\mathbf{A}, \pi \models \exists v \phi$, then there is some $x \in A$ such that $\mathbf{A}, \pi\{v := x\} \models \phi$; the induction hypothesis applied to $\rho = \pi\{v := x\}$ gives $\mathbf{A}, \pi\{v := x\} \models \psi$; and since v does not occur free in ψ (by the condition on the rule) and π agrees with $\pi\{v := x\}$ on all other variables, we have the required $\mathbf{A}, \pi \models \psi$ by (2) of the Compositionality Theorem 3G.1. ⊣

7C. The Completeness Theorem for LPCI. There are actually two versions of the converse of (7-1), both of them useful, and neither of them as plausible as the Soundness Theorem 7B.1 without some work. We give (almost) full proofs of both, starting with some preliminary facts.

7C.1. **Lemma** (Constant Substitution Lemma). *Suppose T is a theory, the variable v does not occur bound in the sequence of formulas*

$$\phi_0, \dots, \phi_n$$

of LPCI(τ)*, and c is a fresh constant, i.e., a constant which does not occur in T or in any ϕ_i; then*

$$\phi_0, \dots, \phi_n \text{ is a deduction from } T$$
$$\iff \phi_0\{v :\equiv c\}, \dots, \phi_n\{v :\equiv c\} \text{ is a deduction from } T.$$

Proof is easy, by taking cases on the justification which allows the insertion of each ϕ_i and the corresponding $\phi_i\{v :\equiv c\}$ in these deductions. The relevant observation (for the direction $\Longrightarrow$ of the equivalence) is that v cannot be the "active variable" in any application of Generalization or Exists Elimination, because if it were, it would occur bound in some formula of the given deduction. ⊣

7C.2. **Lemma** (The Deduction Theorem). *For every theory T, every sentence χ and every formula ϕ,*

$$\text{if } T, \chi \vdash \phi, \text{ then } T \vdash \chi \to \phi.$$

Proof. We follow the architecture of the proof of the Deduction Theorem for the Propositional Calculus, and there are only two additional cases to consider, when χ_n is derived from some earlier formula in the given deduction (from T, χ) by one of the predicate rules of inference.

(13) Generalization. Now $\chi_n \equiv \forall v\psi$, and the induction hypothesis gives us a deduction from T of $\chi \to \psi$; we follow this proof by the formulas

$$\forall v(\chi \to \psi) \text{ (13)}, \ \forall v(\chi \to \psi) \to (\chi \to \forall v\psi) \text{ (10)},$$
$$\chi \to \forall v\psi \text{ (Modus Ponens)}$$

where $\forall v(\chi \to \psi) \to (\chi \to \forall v\psi)$ is an instance of Axiom (10), justified because v does not occur free in χ (which is a sentence).

(14) Exists Elimination. Now $\chi_n \equiv (\exists v\phi \to \psi)$ and it is inferred in the given deduction from T, χ by applying (14) to $\phi \to \psi$, so that v does not occur free in ψ. The induction hypothesis gives us a proof from T of

$$\chi \to (\phi \to \psi),$$

and we follow this with the following sequence of formulas to complete a deduction of $\chi \to (\exists v\phi \to \psi)$ from T:

$$\ldots \phi \to (\chi \to \psi) \text{ (Prop. steps)}$$
$$\exists v\phi \to (\chi \to \psi) \text{ (14)}$$
$$\ldots \chi \to (\exists v\phi \to \psi) \text{ (Prop. steps)},$$

where the application of (14) is justified because v does not occur in $\chi \to \psi$, by the hypothesis and the assumption that χ is a sentence. ⊣

7C.3. **Lemma** (Alphabetic change of bound variables). *If the variable w is free for v in ϕ, then*

$$\vdash \exists v\phi \leftrightarrow \exists w\phi\{v :\equiv w\}$$
$$\vdash \forall v\phi \leftrightarrow \forall w\phi\{v :\equiv w\}$$

Proof. It is enough to show that

$$\vdash \exists w\phi\{v :\equiv w\} \to \exists v\phi,$$

since the rest follows by symmetry, the equivalence of $\forall$ with $\neg\exists\neg$ and propositional logic. Here is a deduction of it:

$$\ldots \phi\{v :\equiv w\} \to \phi\{v :\equiv w\} \quad \text{(Propositional)},$$
$$\phi\{v :\equiv w\} \to \exists v\phi \quad \text{(with } t :\equiv w \text{ in Axiom (11))},$$
$$\exists w\phi\{v :\equiv w\} \to \exists v\phi \quad \text{(Rule (14))}$$

⊣

7C.4. **Lemma** (The natural introduction rules for LPCI). *If T is a set of sentences, the indicated substitutions are free, and the indicated restrictions are obeyed, then the following hold:*

($\to$) *If $T, \chi \vdash \phi$, then $T \vdash \chi \to \phi$.* *Restriction: χ is a sentence.*
($\wedge$) *If $T \vdash \phi$ and $T \vdash \psi$, then $T \vdash \phi \wedge \psi$.*
($\vee$) *If $T \vdash \phi$ or $T \vdash \psi$, then $T \vdash \phi \vee \psi$.*
($\neg$) *If $T, \chi \vdash \psi$ and $T, \chi \vdash \neg\psi$, then $T \vdash \neg\chi$.* *Restriction: χ a sentence.*
($\forall$) *If $T \vdash \phi$, then $T \vdash \forall v\phi$.*
($\exists$) *If $T \vdash \phi\{v :\equiv t\}$, then $T \vdash \exists v\phi$.*

7C.5. **Lemma** (The natural elimination rules for LPCI). *If T is a set of sentences, the indicated substitutions are free, and the indicated restrictions are obeyed, then the following hold:*

($\to$) *If $T \vdash \phi$ and $T \vdash \phi \to \psi$, then $T \vdash \psi$.*
($\wedge$) *If $T \vdash \phi \wedge \psi$, then $T \vdash \phi$ and $T \vdash \psi$.*
($\vee$) *If $T, \phi \vdash \chi$ and $T, \psi \vdash \chi$, then $T, \phi \vee \psi \vdash \chi$*
Restriction: ϕ, ψ are sentences.
($\neg$) *If $T \vdash \neg\neg\phi$, then $T \vdash \phi$.*
($\forall$) *If $T \vdash \forall v\phi$, then $T \vdash \phi\{v :\equiv t\}$*

(∃) *If* $T, \phi\{v :\equiv c\} \vdash \chi$, *then* $T, \exists v\phi \vdash \chi$
Restriction: $\exists v\phi$ *is a sentence and* c *is a constant which does not occur in* T *or in* χ.

Proof. We verify only the last rule of (∃)-elimination whose proof is the fussiest.

By appealing to the Deduction Theorem and Modus Ponens, it is enough to show that

$$\text{if } T \vdash \phi\{v :\equiv c\} \to \chi, \text{ then } T \vdash \exists v\phi \to \chi.$$

So suppose we are given a deduction from T

$$\psi_0, \dots, \psi_n, \phi\{v :\equiv c\} \to \chi$$

in which the constant c may occur, but also the variable v may occur in some ψ_i—which is what causes some complexity in the argument. Choose a variable w which does not occur in any of the formulas in this deduction and notice that the sequence

$$\psi_0\{c :\equiv w\}, \dots, \psi_n\{c :\equiv w\}, \phi\{v :\equiv w\} \to \chi \tag{7-3}$$

is also a deduction from T, where each $\psi_i\{c :\equiv w\}$ is naturally defined by replacing every occurrence of c in ψ_i by the fresh variable w; this follows by the Constant Substitution Lemma 7C.1, since, obviously,

$$\psi_i\{c :\equiv w\}\{w :\equiv c\} \equiv \psi_i,$$

$\phi\{v :\equiv w\}\{w :\equiv c\} \equiv \phi\{v :\equiv c\}$ and c does not occur in χ. We now append to the deduction (7-3) the formula

$$\exists w\phi\{v :\equiv w\} \to \chi \text{ by Rule (14)}$$

to get $T \vdash \exists w\phi\{v :\equiv w\} \to \chi$, which implies $T \vdash \exists v\phi \to \chi$ by Lemma 7C.3 and propositional steps. ⊣

7C.6. **Proposition.** *If* $T \vdash \phi$ *and* $T \vdash \neg\phi$, *then for every sentence* χ, $T \vdash \chi$, *so that, in particular,* $T \vdash \mathbf{ff}$.

Proof. For any sentence χ, by the hypothesis,

$$T, \neg\chi \vdash \phi, \qquad T, \neg\chi \vdash \neg\phi;$$

hence $T \vdash \neg\neg\chi$ by (¬)-introduction, hence $T \vdash \chi$, by (¬)-elimination. ⊣

At this point we can state the first—and most natural—version of the Completeness Theorem.

7C.7. **Theorem** (Gödel's Completeness Theorem, I). *If* τ *is finite, then every logical consequence of a* τ*-theory* T *is a theorem of* T, *i.e., for every sentence* χ,

$$\textit{if } T \models \chi, \textit{ then } T \vdash \chi.$$

The proof will take up the rest of this section and we will need to introduce some new notions and prove various facts about them.

7C.8. **Definition.** A theory T is **consistent** if it does not prove a contradiction, equivalently if $T \nvdash \mathbf{ff}$; it is **inconsistent** in the opposite case, i.e., if $T \vdash \mathbf{ff}$.

One of the basic properties of consistency is that it is *finitely based*:

7C.9. **Lemma.** *A theory T is consistent if and only if every finite subset $T_0 \subseteq T$ of it is consistent.*

Proof is simple and we leave it for Problem x2.42. ⊣

At the same time, if T has a model, then it is consistent, since it cannot be that $\mathbf{A} \models \chi$ for every χ. This observation suggests a different version of the converse to Theorem 7B.1:

7C.10. **Theorem** (Gödel's Completeness Theorem, II). *If τ is finite, then every consistent τ-theory T has a model.*

Proof of CT I from CT II. Suppose $T \models \chi$ but (towards a contradiction), $T \nvdash \chi$. Now $T \cup \{\neg\chi\}$ is consistent, since the opposite assumption gives $T, \neg\chi \vdash \chi$ and then (easily) $T \vdash \chi$. By the assumed version II of the Completeness Theorem, this implies that $T \cup \{\neg\chi\}$ has a model $\mathbf{A}$; but then $\mathbf{A} \models T$ and $\mathbf{A} \not\models \chi$, which contradicts the hypothesis. ⊣

It is equally easy to prove CT II assuming CT I, cf. Problem x2.43, so both versions of the Completeness Theorem express the same, basic mathematical fact.

We now start on the proof of Theorem 7C.10 with a few properties of consistent theories.

7C.11. **Lemma.** *Suppose T is a consistent theory in $\mathsf{LPCI}(\tau)$.*

(1) *For every sentence χ, either $T \cup \{\chi\}$ is consistent, or $T \cup \{\neg\chi\}$ is consistent.*

(2) *If $\exists v\phi$ is a sentence, $T \cup \{\exists v\phi\}$ is consistent and c is a constant which does not occur in T or in $\exists v\phi$, then $T \cup \{\phi\{v :\equiv c\}\}$ is consistent.*

Proof. (1) If both $T \cup \{\chi\}$ and $T \cup \{\neg\chi\}$ are inconsistent, then for any sentence ϕ,

$$T, \chi \vdash \phi \text{ and } T, \neg\chi \vdash \phi,$$

and so by ($\vee$)-elimination $T, \chi \vee \neg\chi \vdash \phi$; but $\vdash \chi \vee \neg\chi$ (propositionally), and so $T \vdash \phi$. Thus T proves every sentence ϕ and it is inconsistent, contrary to hypothesis.

(2) If $T \cup \{\phi\{v :\equiv c\}\}$ is inconsistent, then $T, \phi\{v :\equiv c\} \vdash \psi$ for every sentence ψ; but then, by ($\exists$)-elimination, $T, \exists v\phi \vdash \psi$, and so $T \cup \{\exists v\phi\}$ is inconsistent, contrary to hypothesis. ⊣

7C.12. **Definition.** A theory T in $\mathsf{LPCI}(\tau)$ is **complete** if for each sentence θ of $\mathsf{LPCI}(\tau)$, either $T \vdash \theta$ or $T \vdash \neg\theta$.

The standard example of a complete theory is the theory $\mathrm{Th}(\mathbf{A})$ of a structure, since for every sentence θ, either $\mathbf{A} \models \theta$ or $\mathbf{A} \models \neg\theta$.

7C.13. **Definition.** A theory H is a **Henkin set** in $\mathsf{LPCI}(\tau)$ if it satisfies the following conditions:

(H1) H is consistent.

(H2) For every sentence χ, either $\chi \in H$ or $\neg\chi \in H$, so that in particular, H is complete.

(H3) If $\exists v\phi \in H$, then there is a constant c such that $\phi\{v :\equiv c\} \in H$.

The constant c in the last condition is called a **Henkin witness** for the existential sentence $\exists v\phi$, so briefly:

a Henkin set is a consistent, strongly complete τ-theory which has Henkin witnesses.

7C.14. **Lemma** (Properties of Henkin sets). *Fix a Henkin set H.*

(1) *H is deductively closed, i.e., for every sentence χ,*

$$\text{if } H \vdash \chi, \text{ then } \chi \in H.$$

(2) *For all sentences $\phi, \psi, \exists v\phi, \forall v\phi$:*

$$\begin{aligned}
\neg\phi \in H &\iff \phi \notin H\\
\phi \wedge \psi \in H &\iff \phi \in H \text{ and } \psi \in H\\
\phi \vee \psi \in H &\iff \phi \in H \text{ or } \psi \in H\\
\phi \to \psi \in H &\iff \phi \notin H \text{ or } \psi \in H\\
\exists v\phi \in H &\iff \text{ there is some } c \text{ such that } \phi\{v :\equiv c\} \in H\\
\forall v\phi \in H &\iff \text{ for all } c,\ \phi\{v :\equiv c\} \in H
\end{aligned}$$

PROOF. (1) Suppose $H \vdash \chi$ but $\chi \notin H$; then $\neg\chi \in H$ by (H2), and so $H \vdash \neg\chi$, which makes H inconsistent contradicting property (H1).

(2) We consider just two of these equivalences.

If $\phi \wedge \psi \in H$, then $\phi, \psi \in H$ by the deductive completeness of H, since $\phi \wedge \psi \vdash \phi$ and $\phi \wedge \psi \vdash \psi$; and for the converse of this, we use the fact that $\phi, \psi \vdash \phi \wedge \psi$, so that if $\phi, \psi \in H$, then $H \vdash \phi \wedge \psi$ and so $\phi \wedge \psi \in H$.

If $\exists v\phi \in H$, then $\phi\{v :\equiv c\} \in H$ for some c, by the key property (H3). The converse holds because $\phi\{v :\equiv c\} \vdash \exists v\phi$ and H is deductively complete. ⊣

This lemma suggests that every Henkin set is $\mathrm{Th}(\mathbf{A})$ for some structure $\mathbf{A}$, and so to construct a model of some consistent theory T we should aim to construct a Henkin set which extends T. There are but two, small subtleties that we need to deal with to turn this idea into a proof.

The Completeness Theorem 7C.10 will follow from the following two, crucial lemmas, which have independent interest:

7C.15. **Lemma A.** *Suppose τ is a finite vocabulary.*

Every consistent τ-theory T is contained in a Henkin set $H \supseteq T$ of $\mathsf{LPCI}(\tau^*)$, *where the vocabulary τ^* is an expansion of τ by an infinite sequence of fresh constants $(d_0, d_1, \dots)$, so that*

(7-4) $$\textit{if } \tau = (\mathrm{Const}, \mathrm{Rel}, \mathrm{Funct}, \mathrm{arity}),$$
$$\textit{then } \tau^* = (\mathrm{Const} \cup \{d_0, d_1, \dots\}, \mathrm{Rel}, \mathrm{Funct}, \mathrm{arity}).$$

7C.16. **Lemma B.** *If H is a Henkin set in* $\mathsf{LPCI}(\tau^*)$ (as in Lemma A), *then there is a* $\mathsf{LPCI}(\tau^*)$*-structure*

$$\mathbf{A}^* = (A, \{\overline{c}\}_{c \in \mathrm{Const}}, \{\bar{d}_i\}_{i \in \mathbb{N}}, \{\overline{R}\}_{R \in \mathrm{Rel}}, \{\overline{f}\}_{f \in \mathrm{Funct}}),$$

such that for every $\mathsf{LPCI}(\tau^*)$*-sentence χ,*

$$\mathbf{A}^* \models \chi \iff \chi \in H.$$

In particular, $\mathbf{A}^*$ is a model of H.

Assuming these two lemmas:

Proof of Theorem 7C.10. Suppose T is a consistent $\mathsf{LPCI}(\tau)$-theory, H is a Henkin extension of it in $\mathsf{LPCI}(\tau^*)$ by Lemma 7C.15, and $\mathbf{A}^*$ is a model of H by Lemma 7C.16: now the reduct

$$\mathbf{A} = (A, \{\overline{c}\}_{c \in \mathrm{Const}}, \{\overline{R}\}_{R \in \mathrm{Rel}}, \{\overline{f}\}_{f \in \mathrm{Funct}})$$

is a model of T by (1) of the Compositionality Theorem 3G.1. ⊣

We should mention that the restriction to finite τ in these lemmas is not necessary: in fact they hold for arbitrary τ. But all the interesting applications are to τ-theories with finite τ; and if τ is infinite, then the construction of the (necessarily infinite) τ^* and the proof that $\mathsf{LPCI}(\tau^*)$ has the necessary properties requires some non-trivial results from set theory which are not relevant to the task at hand.

Proof of Lemma A, 7C.15. We assume the hypotheses and the notation of Lemma A.

Sublemma 1. There is a sequence of $\mathsf{LPCI}(\tau^*)$ *sentences*

$$\chi_0, \chi_1, \dots$$

which contains all the $\mathsf{LPCI}(\tau^*)$ *sentences, such that for each n, the constant d_n does not occur in any of the first n sentences $\chi_0, \dots, \chi_{n-1}$.*

Proof. For each $n = 0, 1, \dots,$ let

$$S_n = \text{the set of all sentences of } \mathsf{LPCI}(\tau^*) \text{ of length} \leq 5 + n$$
$$\text{whose variables are in } \{\mathrm{v}_0, \dots, \mathrm{v}_n\}, \text{ and in which}$$
$$\text{only } d_0, \dots, d_{n-1} \text{ of the fresh constants may occur.}$$

The choice of 5 in this definition insures that S_0 is not empty, since

$$\exists \mathrm{v}_0 \mathrm{v}_0 = \mathrm{v}_0, \forall \mathrm{v}_0 = \mathrm{v}_0 \in S_0.$$

At the same time, easily:

1. Each S_n is finite.
2. $S_n \subseteq S_{n+1}$, for each n.
3. d_0 does not occur in any sentence in S_0, and for $n > 0$, d_n does not occur in any sentence of S_{n-1}.

We now enumerate in some standard way all these finite sets,

$$S_n = (\chi_0^n, \dots, \chi_{k_n}^n),$$

and conclude that the required enumeration of all the $\mathsf{LPCI}(\tau^*)$- sentences is the "concatenation" of all these enumerations,

$$\chi_0^0, \dots, \chi_{k_0}^0, \chi_0^1, \dots, \chi_{k_1}^1, \dots . \qquad \dashv \text{(Sublemma 1)}$$

Sublemma 2. There exists a sequence

$$\phi_0, \phi_1, \dots , \tag{7-5}$$

of $\mathsf{LPCI}(\tau^*)$*-sentences with the following properties:*

1. *For each* n, $\phi_{2n} \equiv \chi_n$ *or* $\phi_{2n} \equiv \neg\chi_n$.
2. *For each* n, *if* $\phi_{2n} \equiv \exists v \psi(v)$ *for some variable* v *and formula* $\psi(v)$, *then* $\phi_{2n+1} \equiv \psi(d_n)$, *otherwise* $\phi_{2n+1} \equiv \phi_{2n}$.
3. *For each* n, *the set* $T \cup \{\phi_0, \dots , \phi_{2n+1}\}$ *is consistent.*

Proof. The sentences ϕ_{2n}, ϕ_{2n+1} are defined by recursion on n, using Lemma 7C.11—and their definition is basically determined by the conditions they are required to satisfy. $\dashv$ (Sublemma 2)

It is now easy to verify that the range $H = \{\phi_0, \phi_1, \dots ,\}$ of the sequence of sentences in (7-5) constructed in the proof of Sublemma 2 is a Henkin set.

To see that it includes T, suppose $\chi \in T$. Now $\chi \equiv \chi_n$ for some n, and so either $\phi_{2n} \equiv \chi$ or $\phi_{2n} \equiv \neg\chi$; but $T \cup \{\phi_0, \dots , \phi_{2n+1}\}$ is consistent, and so it cannot contain both χ and $\neg\chi$—so it must be that $\phi_{2n} \equiv \chi$. $\dashv$

Proof of Lemma B, 7C.16. Let H be the Henkin set guaranteed by Lemma A, and let

$$C = \mathrm{Const} \cup \{d_0, d_1, \dots ,\}$$

be the set of the constants in τ^*, including all the fresh constants we added to the vocabulary τ. Lemma 7C.14 suggests that we can construct a model $\mathbf{C}$ of H on the universe C, by setting for $e_1, \dots , e_n, e \in C$,

$$R^{\mathbf{C}}(e_1, \dots , e_n) \iff R(e_1, \dots , e_n) \in H,$$
$$f^{\mathbf{C}}(e_1, \dots , e_n) = e \iff f(e_1, \dots , e_n) = e \in H,$$

and this almost works, except that it gives "multiple valued" interpretations of the constants: it may well be that $c = c' \in H$ for two distinct constants in τ^*. To deal with this, we need to "identify" constants which H thinks are equal, as follows. We set:

$$a \sim b \iff (a = b) \in H \qquad (a, b \in C).$$

Sublemma 1. The relation $\sim$ is an equivalence relation on the set C of constants, i.e., for all $a, b, c \in C$:

$$a \sim a, \quad a \sim b \Longrightarrow b \sim a, \quad [a \sim b \ \&\ b \sim c] \Longrightarrow a \sim c$$

Proof is immediate from the axioms of equality, which are satisfied by H, since it is deductively closed. For example, $(a = a) \in H$ for every constant a, simply because $\vdash a = a$. ⊣ (Sublemma 1)

The same kind of argument gives the next two facts we need:

Sublemma 2. For every n-ary relation symbol R and for all constants $a_1, \ldots, a_n, b_1, \ldots, b_n \in C$,

$$a_1 \sim b_1, \ldots a_n \sim b_n \Longrightarrow [R(a_1, \ldots, a_n) \in H \iff R(b_1, \ldots, b_n) \in H].$$

Sublemma 3. For every n-ary function symbol f and all constants $a_1, \ldots, a_n \in C$, there is a constant $u \in C$ such that

$$(f(a_1, \ldots, a_n) = u) \in H;$$

and for all constants $a_1, \ldots, a_n, u, b_1, \ldots, b_n, v \in C$,

$$a_1 \sim b_1, \ldots a_n \sim b_n, u \sim v$$
$$\Longrightarrow (f(a_1, \ldots, a_n) = u) \in H \iff (f(b_1, \ldots, b_n) = v) \in H.$$

We now let $\widetilde{C}$ be the set of equivalence classes of this equivalence relation on C and choose a set $A \subseteq C$ of representatives for them: this means that we define a function $a \mapsto \overline{a}$ on C to itself such that

$$a \sim b \iff \overline{a} = \overline{b}.$$

This set A is the universe of the structure $\mathbf{A}^*$ that we want to construct. On it we interpret each constant c by $\overline{c}$, and for the relation and function symbols as set

$$\overline{R}(\overline{a}_1, \ldots, \overline{a}_n) \iff R(a_1, \ldots, a_n) \in H,$$
$$\overline{f}(\overline{a}_1, \ldots, \overline{a}_n) = \overline{e} \iff (f(a_1, \ldots, a_n) = e) \in H.$$

The Sublemmas insure that these are *good definitions*; and then Lemma 7C.14 implies easily that for this structure $\mathbf{A}$ and all χ in $\mathsf{LPCI}(\tau^*)$,

$$\mathbf{A} \models \chi \iff \chi \in H,$$

so that, in particular, $\mathbf{A}$ is a model of H. ⊣

7D. The Compactness and Skolem-Löwenheim Theorems. We end the section with two important corollaries of the Completeness Theorem which are formulated entirely in semantic terms, i.e., without reference to the proof theory of LPCI.

7D.1. **Theorem** (The Compactness Theorem). *If every finite subset of a theory T in a finite vocabulary has a model, then T has a model.*

7D.2. **Theorem** (The Weak Skolem-Löwenheim Theorem). *If a theory T in a finite vocabulary has a model, then it has a model $\mathbf{A} = (A, \dots)$ whose universe is countable, i.e., $A = \{a_0, a_1, \dots ,\}$ is the range of a sequence.*

We will leave the (easy now) proofs of these basic results and some of their consequences for the problems, except for the following, very interesting application of compactness:

7E. Non-standard models. Let $n \mapsto \Delta(n)$ be the function from natural numbers to terms of the language of arithmetic defined by the recursion,

$$\Delta(0) :\equiv 0, \quad \Delta(n+1) :\equiv S(\Delta(n)), \tag{7-6}$$

so that $\Delta(1) \equiv S(0)$, $\Delta(2) \equiv S(S(0))$, etc. These **numerals** are the standard (unary) names of numbers in the language of arithmetic.

Let $\tau(c) = (0, c, S, +, \cdot)$ be the expansion of the vocabulary of arithmetic by a new constant c and let

$$T(c) = \mathrm{Th}(\mathbf{N}) \cup \{c \neq \Delta(0), c \neq \Delta(1), \dots\}. \tag{7-7}$$

It is easy to check that $T(c)$ is consistent (Problem x2.46), and so by Theorem 7C.10 it has a model

$$\mathbf{A} = (A, \overline{0}, \overline{c}, \overline{S}, \overline{+}, \overline{\cdot}).$$

This model satisfies all the true sentences in the language of arithmetic, since $\mathrm{Th}(\mathbf{N}) \subseteq T(c)$, and so its reduct

$$\mathbf{N}^* = (A, \overline{0}, \overline{S}, \overline{+}, \overline{\cdot}) \tag{7-8}$$

to the language of arithmetic satisfies all the sentences which are true in $\mathbf{N}$. But $\mathbf{N}^*$ is not isomorphic with $\mathbf{N}$: because if $\pi : \mathbb{N} \rightarrowtail\!\!\!\!\rightarrow A$ were an isomorphism, then (easily)

$$\pi(n) = \Delta(n)^{\mathbf{A}},$$

and the interpretations of the numerals do not exhaust the universe A since $\mathbf{A} \models c \neq \Delta(n)$ for every n.

In short, $\mathbf{N}^*$ is a structure which is elementarily equivalent (as on page 26) with $\mathbf{N}$ but not isomorphic with $\mathbf{N}$: it is a **non-standard model of true arithmetic** and it has a very interesting structure.

§8. Problems.

x2.1. Give rigorous definitions of the $\mathsf{LPCI}(\tau)$ terms and formulas.

x2.2. Determine the free and bound occurrences of variables in the following (misspelled) formula of $\mathsf{LPCI}(\leq)$:

$$\phi \equiv \forall y(x \leq y) \wedge \forall x \exists y(x \leq y \wedge \neg(y \leq x))$$

Which are the free variables of ϕ and which are its bound variables?

x2.3. Consider the following two sentences in the language of posets:

$$\phi \equiv \exists \mathrm{v}_1 \exists \mathrm{v}_2 \forall \mathrm{v}_2 [\mathrm{v}_1 \leq \mathrm{v}_2], \quad \psi \equiv \exists \mathrm{v}_1 \forall \mathrm{v}_2 \exists \mathrm{v}_2 [\mathrm{v}_1 \leq \mathrm{v}_2].$$

What do they mean, and do they have the same truth value in every poset?

x2.4. In the language $(\leq)$ of posets first abbreviate

$$x < y :\equiv x \leq y \wedge \neg(x = y)$$

and consider the following sentence χ which says that "$(P, \leq)$ has a minimal element",

$$\chi \equiv (\exists x)(\forall y)(y \not< x).$$

Write out the correctly spelled form of χ (with formal variables $\mathrm{v}_0, \mathrm{v}_1$, the correct spelling for prime formulas and all the parentheses).

x2.5. Write out the correctly spelled form of $(\exists! x)\phi$ in (2-2).

x2.6. Give an example of a term α, variables v_1, v_2 and terms t_1, t_2 such that $\alpha\{v_1 :\equiv t_1\}\{v_2 :\equiv t_2\} \not\equiv \alpha\{v_1 :\equiv t_1, v_2 :\equiv t_2\}$.

x2.7. Prove that the function $\sigma(x) = x + 1$ is an automorphism of the usual linear ordering $(\mathbb{Q}, \leq)$ on the rational numbers.

x2.8. Prove that for every structure $\mathbf{A}$, the identity $\sigma(x) = x$ on A is an automorphism—the *trivial* one. Prove also that the structure $\mathbf{N}$ of arithmetic has no other automorphisms—it is *rigid*.

x2.9. With the notation for simultaneous updates in Section 3F, prove that

$$\pi\{v_1 := x_1, \ldots, v_n := x_n\} = \pi\{v_1 := x_1\}\{v_2 := x_2\} \cdots \{v_n := x_n\}.$$

x2.10. Prove Theorem 3F.1, the Tarski conditions for the satisfaction relation.

x2.11. For each formula χ of the Propositional Calculus PL, any sequence $p_1, \ldots, p_n$ of distinct propositional variables which includes all the variables which occur in χ and any sequence $\phi_1, \dots, \phi_n$ of $\mathsf{LPCI}(\tau)$-formulas, let

$$\chi^* \equiv \chi\{p_1 :\equiv \phi_1, \dots, p_n :\equiv \phi_n\}$$
$$\equiv \text{ the result of replacing each } p_i \text{ in } \chi \text{ by } \phi_i.$$

Fix a τ-structure $\mathbf{A}$ and for any assignment π into $\mathbf{A}$ let v_π be any truth-value assignment to the propositional variables such that

$$v_\pi(p_i) = \begin{cases} 1, & \text{if } \mathbf{A}, \pi \models \phi, \\ 0, & \text{otherwise.} \end{cases}$$

Prove that $v_\pi \models \chi \iff \mathbf{A}, \pi \models \chi^*$ and infer that

$$\text{if } \chi \text{ is a tautology, then for every } \pi,\ \mathbf{A}, \pi \models \chi^*.$$

x2.12. Prove that if $\sigma : \mathbf{A} \rightarrowtail\!\!\!\!\twoheadrightarrow \mathbf{B}$ is an isomorphism, then its inverse $\sigma^{-1} : \mathbf{B} \rightarrowtail\!\!\!\!\twoheadrightarrow \mathbf{A}$ is also an isomorphism. Infer that isomorphism is an *equivalence condition* between structures, i.e., for all $\mathbf{A}, \mathbf{B}, \mathbf{C}$,

$$\mathbf{A} \cong \mathbf{A}, \quad \mathbf{A} \cong \mathbf{B} \Longrightarrow \mathbf{B} \cong \mathbf{A}, \quad [\mathbf{A} \cong \mathbf{B} \wedge \mathbf{B} \cong \mathbf{C}] \Longrightarrow \mathbf{A} \cong \mathbf{C}.$$

x2.13. Suppose $\sigma : A \rightarrowtail\!\!\!\!\twoheadrightarrow B$ is an isomorphism between the two τ-structures $\mathbf{A}, \mathbf{B}$.

(1) Prove that for every extended term $t(\vec{v})$ and all $x_1, \dots, x_n \in A$,

$$\sigma(t^{\mathbf{A}}[x_1, \dots ,_n]) = t^{\mathbf{B}}[\sigma(x_1), \dots, \sigma(x_n)].$$

(2) Prove that for every extended formula $\chi(\vec{v})$ and all $x_1, \dots, x_n \in A$,

$$\chi^{\mathbf{A}}[x_1, \dots, x_n] \iff \chi^{\mathbf{B}}[\sigma(x_1), \dots, \sigma(x_n)].$$

(3) Infer that if $\sigma : A \rightarrowtail\!\!\!\!\twoheadrightarrow A$ is an automorphism of $\mathbf{A}$ and $P(x_1, \dots, x_n)$ is $\mathbf{A}$-elementary, then

$$P(x_1, \dots, x_n) \iff P(\sigma(x_1), \dots, \sigma(x_n)).$$

x2.14. Prove that if a binary relation $P(x, y)$ is elementary in a structure $\mathbf{A}$, then so is the *converse relation*

$$\breve{P}(x, y) \iff P(y, x).$$

x2.15. Prove that if $f(\vec{x}), g(\vec{x})$ are elementary functions in a structure $\mathbf{A}$, then so is the relation

$$P(\vec{x}) \iff f(\vec{x}) = g(\vec{x}).$$

In the next two problems you are asked to decide whether a given relation is elementary or not on a given structure and to provide an extended formula which defines it if your answer is “yes”. For example, the relation

$$x \mid y \iff x \text{ divides } y$$

is elementary on the structure of arithmetic **N** (arithmetical), defined by the extended formula

$$\phi(x, y) \equiv \exists z(y = x \cdot z).$$

You will not be able to prove your negative answers, as we have not developed yet tools for proving non-elementarity, but you should try to guess the correct answers.

x2.16*. Determine whether the following relations are elementary on a fixed, symmetric graph $\mathbf{G} = (G, E)$, and if your answer is positive, find a formula which defines them.

(1) $P(x, y) \iff d(x, y) \leq 2$.
(2) $P(x, y) \iff d(x, y) = 2$.
(3) $P(x, y) \iff d(x, y) < \infty$.
(4) $P(x, y, z) \iff d(x, y) \leq d(x, z)$
(5) $P(x) \iff$ every y can be joined to x.

x2.17. Determine whether the (usual) ordering relation on real numbers is elementary on the field $\mathbf{R} = (\mathbb{R}, 0, 1, +, \cdot)$, and if your answer is positive, find a formula which defines them. (You need to know something about the real numbers to do this.)

x2.18* (Dedekind's characterization of **N**). Prove that every two structures

$$\mathbf{N}_1 = (\mathbb{N}_1, 0_1, S_1, +_1, \cdot_1), \quad \mathbf{N}_2 = (\mathbb{N}_2, 0_2, S_2, +_2, \cdot_2)$$

which satisfy the Peano axioms in 1D are isomorphic. Hint: Let

$$X = \Big\{ t \in \mathbb{N}_1 \mid \text{ there exists a function } f : A \to B \text{ such that}$$
$$0_1 \in A, f(0_1) = 0_2 \in \mathbb{N}_2, \text{ and for all } t \in \mathbb{N}_1,$$
$$S_1(t) \in A \Longrightarrow \Big[t \in A \;\&\; f(t) \in \mathbb{N}_2 \;\&\; f(S_1(t)) = S_2(f(t)) \Big] \Big\}.$$

Use the Induction Axiom on $\mathbf{N}_1$ to prove that $X = \mathbb{N}_1$ and there is an injection $\sigma : \mathbb{N}_1 \rightarrowtail \mathbb{N}_2$ such that

$$\sigma(0_1) = 0_2 \text{ and for all } t \in \mathbb{N}_1, \sigma(S_1(t)) = S_2(\sigma(t));$$

and then use the Induction Axiom on $\mathbf{N}_2$ to prove that $\sigma[N_1] = N_2$.

x2.19 (The Division Theorem for **N**). Prove that for every $x \in \mathbb{N}$ and every $y > 0$ in $\mathbb{N}$, there exist unique $q, r \in \mathbb{N}$ such that

$$x = y \cdot q + r, \quad 0 \leq r < y.$$

We set $\text{quot}(x, y) = q$ and $\text{rem}(x, y) = r$, and for completeness, we also let $\text{quot}(x, 0) = 0, \text{rem}(x, 0) = x$.

x2.20. Suppose that $d_0, \ldots, d_t \in \mathbb{N}$ are *relatively prime*, i.e., no two of them have a common factor > 1; prove that if $x \in \mathbb{N}$ and $d_i \mid x$ for $i = 0, \ldots, d_t$, then $d_0 d_1 \cdots d_t \mid x$.

x2.21. Determine whether the following relations are arithmetical, i.e., elementary on the structure of arithmetic $\mathbf{N} = (\mathbb{N}, 0, S, +, \cdot)$.

1. $\mathrm{Prime}(x) \iff x$ is a prime number.
2. $\mathrm{TP}(x) \iff$ there are infinitely many twin primes y such that $x \leq y$.
3. $\mathrm{Exp}(x, w) \iff 2^x = w$, where 2^x is defined as usual,
$$2^0 = 1, \quad 2^{x+1} = 2 \cdot 2^x.$$
4. $\mathrm{Quot}(x, y, w) \iff \mathrm{quot}(x, y) = w$.
5. $\mathrm{Rem}(x, y, w) \iff \mathrm{rem}(x, y) = w$.
6. $x \perp y \iff x$ and y are coprime (i.e., no number other than 1 divides both x and y).

x2.22. Prove that the following functions and relations on $\mathbb{N}$ are arithmetical.

1. $p(i) = p_i =$ the i'th prime number, so that $p_0 = 2, p_1 = 3, p_2 = 5$, etc.
2. $f_n(x_0, \ldots, x_n) = p_0^{x_0+1} \cdot p_1^{x_1+1} \cdots p_n^{x_n+1}$. (This is a different function of $n+1$ arguments for each n.)
3. $R(u) \iff$ there exists some n and some $x_1, \ldots, x_n$ such that $u = f_n(x_1, \ldots, x_n)$.

x2.23. The Ackermann function is defined by the following *double recursion*:

$$A(0, x) = x + 1$$
$$A(n+1, 0) = A(n, 1)$$
$$A(n+1, x+1) = A(n, A(n+1, x))$$

1. Compute $A(1, 2)$.
2. Compute $A(2, 1)$.
3. Prove that the Ackermann function is arithmetical.

x2.24. Prove Lemma 4B.1.

x2.25. Prove that the ring of integers $\mathbf{Z}$ admits coding of tuples (Example 4B.2).

x2.26. Prove (1) – (3) of Proposition 5A.1. HINT: Use Problem x2.11.

x2.27. Prove (4) and (5) of Proposition 5A.1.

x2.28. Prove (6) of Proposition 5A.1 and infer Corollary 5A.2.

x2.29. (1) Let $\mathbf{L} = ([0,1), 0, \leq)$, where $[0,1)$ is the half-open interval of real numbers,

$$[0,1) = \{x \in \mathbb{R} \mid 0 \leq x < 1\}$$

and 0 is a constant which names the number 0. Prove that $\mathbf{L}$ admits effective elimination of quantifiers. Infer that it is a decidable structure, i.e., there is an effective procedure which decides whether $\mathbf{L} \models \chi$, for an arbitrary sentence χ.

(2) Let $\mathbf{L}' = ([0,1), \leq)$ be the same linear ordering as in (1), but in the language without a name for 0. Does $\mathbf{L}'$ admit elimination of quantifiers?

(3) Is the structure $\mathbf{L}'$ decidable?

In the next three problems you are asked to decide whether a class Φ of τ-structures is basic elementary, elementary or neither, and if your answer is positive for one of these questions to define the relevant theory. For example, when $\tau = (\leq)$ is the signature of partial orderings (with just one binary relation symbol $\leq$),

$$\begin{aligned}&(P, \leq) \text{ is a partial ordering}\\ &\quad\iff (P, \leq) \models (\forall x)[x \leq x] \wedge (\forall x)(\forall y)[(x \leq y \wedge y \leq x) \rightarrow x = y]\\ &\qquad\qquad\qquad \wedge (\forall x)(\forall y)(\forall z)[(x \leq y \wedge y \leq z) \rightarrow x \leq z].\end{aligned}$$

You will not be able to prove your negative answers as we have not developed yet tools for proving that a class of structures is not basic elementary or elementary, but you should try to guess the correct answers.

x2.30. For the empty signature τ (for which the τ-structures are just sets) decide whether the following properties of τ-structures are basic elementary, elementary or neither; and if your answer is positive to one of these, find a theory which axiomatizes the given property:

1. A has exactly 3 elements.
2. A is finite.
3. A is infinite.

x2.31. For the signature $\tau = (\leq)$ with just one binary relation symbol, decide whether the following properties of τ-structures are basic elementary or elementary, and if your answer is positive, define the relevant theory:

1. $(P, \leq)$ is an infinite partial ordering.
2. $(P, \leq)$ is a finite partial ordering.
3. $(P, \leq)$ is an infinite linear ordering.
4. $(P, \leq)$ is a finite linear ordering.

x2.32. For the signature $\tau = (E)$ with just one binary relation symbol, decide whether the following properties of τ-structures (graphs) are basic elementary or elementary, and if your answer is positive, define the relevant theory:

1. (G, E) is a symmetric graph.
2. (G, E) is symmetric and connected.

x2.33*. Prove that isomorphic structures are elementarily equivalent,

$$\mathbf{A} \cong \mathbf{B} \Longrightarrow \mathbf{A} \approx \mathbf{B}.$$

x2.34. Prove that every bijection $\sigma : A \rightarrowtail\!\!\!\!\twoheadrightarrow A$ of a set A with itself is an automorphism of the trivial structure (A) with no primitives. Use this fact to identify all unary and binary (A)-elementary relations.

x2.35. Consider the structure $(\mathbb{Q}, \leq)$ of the rational numbers with their ordering.

1. Find all unary, elementary relations in $(\mathbb{Q}, \leq)$.
2. Find all binary, elementary relations in $(\mathbb{Q}, \leq)$.

x2.36. Construct a model of the Robinson system Q which is not isomorphic with the standard model $\mathbf{N}$. HINT: Take for universe $A = \mathbb{N} \cup \{\infty\}$ for some object $\infty \notin \mathbb{N}$.

x2.37*. Construct a model of the Robinson system Q in which addition is not commutative.

x2.38. Give an example which shows that the restriction is necessary in Axiom scheme (10) for $\mathsf{LPCI}(\tau)$.

x2.39. Give an example which shows that the restriction on the Exists Elimination Rule (14) is necessary.

x2.40. Show that if $T \vdash \forall v \phi(v, \vec{u})$ and x is any variable which is free for v in $\phi(v, \vec{u})$, then $T \vdash \forall x \phi(x, \vec{u})$.

x2.41. Let ϕ be a formula of the Propositional Calculus PL whose variables are included in the list of distinct variables $p_1, \dots, p_n$, and let $\psi_1, \dots, \psi_n$ be LPCI-formulas. The substitution

$$\phi\{p_1 :\equiv \psi_1, \dots, p_m :\equiv \psi_n\} \tag{8-1}$$

is defined in the obvious way, by replacing each p_i in ϕ by ψ_i.

(1) Prove that it ϕ is a propositional tautology, then the formula (8-1) is a propositional theorem of LPCI.

(2) Prove that if χ is a propositional theorem of LPCI, then there is a propositional tautology ψ such that for suitable $p_1, \dots, p_n$ and $\psi_1, \dots, \psi_n$, $\chi \equiv \phi\{p_1 :\equiv \psi_1, \dots, p_m :\equiv \psi_n\}$.

x2.42. Prove Lemma 7C.9, that a theory is consistent if and only if all its finite subsets are consistent.

x2.43. Prove that version I of the Completeness Theorem 7C.7 implies version II, Theorem 7C.10.

x2.44. Prove the Compactness Theorem 7D.1.

x2.45. Prove the Skolem-Löwenheim Theorem 7D.2.

x2.46. Prove that the theory $T(c)$ in (7-7) is consistent. HINT: Show that each finite subset of $T(c)$ has a model.

x2.47. Prove that the structure $\mathbf{N}^*$ defined in (7-8) is not isomorphic with the standard model of arithmetic $\mathbf{N} = (\mathbb{N}, 0, S, +, \cdot)$.

x2.48. Prove that if a τ-theory T has arbitrarily large, finite models, then it has an infinite model.

x2.49. For the empty signature τ (for which the τ-structures are just sets) decide whether the following properties of τ-structures are basic elementary or elementary, and prove your answer.

1. A is finite.
2. A is infinite.

x2.50. For the signature $\tau = (E)$ with just one, binary relation symbol, prove that the class of structures which are symmetric, connected graphs is not elementary.

x2.51. (1) Prove that if a sentence χ in the language $(0, 1, +, \cdot)$ of fields is true in every field $\mathbf{F}$ of characteristic 0, then there is a number p_0 such that χ is true in every field $\mathbf{F}$ of characteristic $p \geq p_0$.

(2) Infer that the theory Fields_0 of fields of characteristic 0 is not a basic elementary class.

x2.52. There is an apparent contradiction between

(1) the construction in the last section 7E of a model $\mathbf{N}^*$ of Peano Arithmetic which is not isomorphic with $\mathbf{N}$, and
(2) Dedekind's characterization of $\mathbf{N}$ in Problem x2.18*.

Discuss and explain.

x2.53*. Let $\mathbf{N}^*$ be a non-standard model of true arithmetic as in Section 7E, i.e., $\mathbf{N}^*$ is elementarily equivalent but not isomorphic with $\mathbf{N}$. Prove that if we define on $\mathbf{N}^*$ the relation

$$xE^*y \iff (x +^* 1 = y) \vee (y +^* 1 = x),$$

then the following two relations (from Problem x2.16*) are not elementary in $\mathbf{N}^*$—and hence not elementary in the graph $(\mathbb{N}^*, E^*)$:

(3) $P(x, y) \iff d(x, y) < \infty$.
(4) $P(x, y, z) \iff d(x, y) \leq d(x, z)$.

HINT: The *standard part* of $\mathbf{N}^*$ is an initial segment of $\mathbf{N}^*$ which is isomorphic with $\mathbf{N}$. We may assume that it is $\mathbf{N}$ and put

$$\mathrm{Inf} = \mathbb{N}^* \setminus \mathbb{N} = \text{ the set of “infinite numbers” in } \mathbb{N}^*.$$

This set is not empty. For (3), prove and use the fact that Inf is not elementary; and for (4) prove and use the stronger fact, that Inf is not *elementary from a parameter*, i.e., for every extended formula $\chi(u, v)$ of arithmetic and every $z \in \mathbb{N}^*$,

$$\mathrm{Inf} \neq \{x \in \mathbb{N}^* \mid \chi^{\mathbf{N}^*}[x, z]\}.$$