Topics in Applied Math: Logic and Foundations of Mathematics

Part 2. First order theory

Kazuyuki Tanaka

BIMSA

October 11, 2025

清華大学 求真书院
Qiuzhen College, Tsinghua University

Logic and Foundations

- **Part 1. Equational theory**
- **Part 2. First order theory**
- **Part 3. Basic Model theory**
- **Part 4. First order arithmetic and incompleteness theorems**
- **Part 5. Models of first-order arithmetic**
- **Part 6. Second order arithmetic and reverse mathematics**

Part 2. Schedule

- Oct. 10, (1) First order logic: formal systems and structures
- Oct. 11, (2) Gödel's completeness theorem and applications

Recap

- First-order logic is developed in the common logical symbols and specific mathematical symbols. Major logical symbols are propositional connectives, quantifiers $\forall x$ and $\exists x$ and equality $=$. The set of mathematical symbols to use is called a **language**.
- A **structure** in language $\mathcal{L}$ (simply, a $\mathcal{L}$-structure) is defined as a non-empty set A equipped with an interpretation of the symbols in $\mathcal{L}$.
- A **term** is a symbol string to denote an element of a structure. A **formula** is a symbol string to describe a property of a structure. A formula without free variables is called a **sentence**.
- "A sentence φ is **true** in $\mathcal{A}$, written as $\mathcal{A} \models \varphi$" is defined by Tarski's clauses. The truth of a formula with free variables is defined by the truth of its universal closure.
- A set of sentences in the language $\mathcal{L}$ is called a **theory**. $\mathcal{A}$ is a **model** of T, denoted by $\mathcal{A} \models T$, if $\forall \varphi \in T \ (\mathcal{A} \models \varphi)$.
- We say that φ holds in T, written as $T \models \varphi$, if $\forall \mathcal{A}(\mathcal{A} \models T \to \mathcal{A} \models \varphi)$.

Definition 2.2 (Gentzen-Tait system $\mathrm{GT}(T)$ of a theory T)

A sequent $\varphi_1, \ldots, \varphi_n$ (i.e., a multi-set of formulas) intuitively means $\varphi_1 \vee \cdots \vee \varphi_n$.
A formula φ is automatically transformed into the negation normal form, i.e., constructed from atomic formulas or their negations by means of $\wedge$, $\vee$, $\forall$, and $\exists$.

Axioms

(0) φ (where $\varphi \in T$)

(1) Law of excluded middle: $\neg\psi, \psi$ (where ψ is an atomic formula)

(2) axioms of equality: (i) $x = x$, (ii) $x \neq y, y = x$, etc.

Inference rules

$$\frac{\Gamma, \varphi, \psi}{\Gamma, \varphi \vee \psi}\ (\vee), \quad \frac{\Gamma, \varphi \quad \Gamma, \psi}{\Gamma, \varphi \wedge \psi}\ (\wedge), \quad \frac{\Gamma, \varphi(t)}{\Gamma, \exists x \varphi(x)}\ (\exists), \quad \frac{\Gamma, \varphi(x)}{\Gamma, \forall x \varphi(x)}\ (\forall)(x \text{ is not free in } \Gamma)$$

$$\frac{\Gamma}{\Delta}\ (\mathrm{weak})(\Gamma \text{ is a subsequence of } \Delta), \quad \frac{\Gamma, \neg A \quad \Gamma, A}{\Gamma}\ (\mathrm{cut})$$

- A **proof tree** in the system GT(T) is a finite tree in which each vertex is labelled with a sequent so that a sequent at each top vertex (leaf) is an axiom, and the sequents of adjacent nodes express an inference rule. See an example below.
- If there is a proof tree rooted at a sequent Γ, we write it as $T \vdash \Gamma$. Such a tree is called a **proof** of $T \vdash \Gamma$ (or a **proof** of Γ in T).
- If $T = \varnothing$ or T is clear from the context, we omit T and write $\vdash \Gamma$.

Example 5

For any term t,

$$\dfrac{\dfrac{\dfrac{x = x}{\forall x(x = x)}\,(\forall)}{\forall x(x = x), t = t}\,(\text{weak}) \qquad \dfrac{t \neq t, t = t}{\exists x(x \neq x), t = t}\,(\exists)}{t = t}\,(\text{cut})$$

Lemma 2.4

$\vdash \neg\varphi, \varphi$ for any formula φ.

Lemma 3.1 (Deduction theorem)

Let T be an $\mathcal{L}$-theory, φ a sentence and Γ be a sequent. Then,

$$T \cup \{\varphi\} \vdash \Gamma \Rightarrow T \vdash \neg\varphi, \Gamma.$$

Proof. By adding $\neg\varphi$ to all the sequents appearing in the proof tree of $T \cup \{\varphi\} \vdash \Gamma$ and modifying it at the top nodes, we can construct a proof tree of $T \vdash \neg\varphi, \Gamma$. □

Definition 3.2

T is said to be **inconsistent** if $T \vdash$ (i.e., T proves the empty sequent). Otherwise, T is said to be **consistent**.

Lemma 3.3

Let T be a theory and φ a sentence. The following hold.

(1) If there exists a φ such that $T \vdash \varphi$ and $T \vdash \neg\varphi$, then T is inconsistent.

(2) $T \cup \{\neg\varphi\}$ is inconsistent $\Leftrightarrow T \vdash \varphi$.

(3) $T \cup \{\neg\varphi\}$ is consistent $\Leftrightarrow T \vdash \varphi$ does not hold .

Lemma 3.4

If T is consistent, $T \cup \{\varphi\}$ or $T \cup \{\neg\varphi\}$ are consistent for any sentence φ.

Proof. It is clear from (1) and (3) in the above lemma.

Lemma 3.5

Let $T \cup \{\exists x \varphi(x)\}$ be consistent and c be a new constant. Then $T \cup \{\varphi(\mathsf{c})\}$ is also consistent in $\mathcal{L}' = \mathcal{L} \cup \{\mathsf{c}\}$.

Proof.

- By way of contradiction, assume $T \cup \{\varphi(\mathsf{c})\} \vdash$. Then, we have $T \vdash \neg\varphi(\mathsf{c})$ by the deduction theorem. Replacing all c with x in its proof, we have a proof of $T \vdash \neg\varphi(x)$.
- Adding an inference rule ($\forall$) at the root of the proof tree, we have a proof of $T \vdash \forall x \neg\varphi(x)$, that is, $T \vdash \neg\exists x \varphi(x)$.
- Therefore, by (2) of lemma in the previou page, $T \cup \{\exists x \varphi(x)\}$ is inconsistent, which contradicts with our assumption.

Lemma 3.6

Let $\mathcal{L}$ be a language. Then there exists a set C of constants not included in $\mathcal{L}$ and a set H of sentences in $\mathcal{L}' = \mathcal{L} \cup C$ such that for any consistent $\mathcal{L}$-theory T, the following hold:

(1) $T \cup H$ is consistent.

(2) For each $\mathcal{L}'$-sentence $\exists x \varphi(x)$ such that $T \cup H \vdash \exists x \varphi(x)$, there exists $\mathsf{c} \in C$ such that $T \cup H \vdash \varphi(\mathsf{c})$.

- For an equational theory, we can construct a model by dividing the term algebra by the congruence relation.
- For first-order logic, the term algebra of the given language is not sufficient to make a model. So, we extend the language by introducing many new constants called the **Henkin constants**.

Proof. We describe how to construct C and H of the lemma. For each $\mathcal{L}$-sentence $\exists x\varphi(x)$, we add a new constant $\mathsf{c}_{\exists x\varphi(x)}$, and collect them as C_1, i.e.,

$$C_1 = \{\mathsf{c}_{\exists x\varphi(x)} : \exists x\varphi(x) \text{ is an } \mathcal{L}\text{-sentence}\}.$$

Then, we define a set H_1 of sentences as follows,

$$H_1 = \{\neg\exists x\varphi(x) \vee \varphi(\mathsf{c}_{\exists x\varphi(x)}) : \exists x\varphi(x) \text{ is an } \mathcal{L}\text{-sentence and } \mathsf{c}_{\exists x\varphi(x)} \in C_1\}$$

By a simple logical calculation, we have $\vdash \exists x(\neg\exists x\varphi(x) \vee \varphi(x))$. Since T is consistent, $T \cup \{\exists x(\neg\exists x\varphi(x) \vee \varphi(x))\}$ is also consistent. By the last lemma, $T \cup H_1$ is consistent. Similarly, for each sentence of the form $\exists x\varphi(x)$ in $\mathcal{L}_1 = \mathcal{L} \cup C_1$, we add a constant $\mathsf{c}_{\exists x\varphi(x)}$ and collect them as $C_2 \supseteq C_1$. We also define $H_2 \supseteq H_1$ as follows.

$$C_2 = \{\mathsf{c}_{\exists x\varphi(x)} : \exists x\varphi(x) \text{ is an } \mathcal{L}_1\text{-}sentence\}.$$

$$H_2 = \{\neg\exists x\varphi(x) \vee \varphi(\mathsf{c}_{\exists x\varphi(x)}) : \mathsf{c}_{\exists x\varphi(x)} \in C_2\}$$

Then, $T \cup H_2$ is consistent. By repeating this process, we construct

$$C_0 = \varnothing \subseteq C_1 \subseteq C_2 \subseteq \cdots, \qquad H_0 = \varnothing \subseteq H_1 \subseteq H_2 \subseteq \cdots$$

and we set

$$C = \bigcup_{i \in \mathbb{N}} C_i \text{ and } H = \bigcup_{i \in \mathbb{N}} H_i.$$

Proof (continued). Next, we show that C and H conditions (1) and (2) of the lemma.

(1) If $T \cup H$ were inconsistent, then the inconsistency results from some finite segment of $T \cup H$, that is, there would exist $i \in \mathbb{N}$ such that $T \cup H_i$ were inconsistent. This contradicts the construction of $\{H_i\}$.

(2) Let $\mathcal{L}' = \mathcal{L} \cup C$ and $\exists x \varphi(x)$ be an $\mathcal{L}'$-sentence. Then $\exists x \varphi(x)$ is an $\mathcal{L} \cup C_i$-sentence for some $i \in \mathbb{N}$. So,

$$\neg \exists x \varphi(x) \vee \varphi(\mathsf{c}_{\exists x \varphi(x)}) \in H_{i+1} \subseteq H.$$

Now, we assume $T \cup H \vdash \exists x \varphi(x)$. Then we have $T \cup H \vdash \exists x \varphi(x), \varphi(\mathsf{c}_{\exists x \varphi(x)})$ by weakening. Since we have $\vdash \neg \varphi(\mathsf{c}_{\exists x \varphi(x)}), \varphi(\mathsf{c}_{\exists x \varphi(x)})$ by the excluded middle, we deduce $T \cup H \vdash \exists x \varphi(x) \wedge \neg \varphi(\mathsf{c}_{\exists x \varphi(x)}), \varphi(\mathsf{c}_{\exists x \varphi(x)})$ by $(\wedge)$. Thus, by inference rule (cut), we have

$$T \cup H \vdash \varphi(\mathsf{c}_{\exists x \varphi(x)}).$$

□

A constant in C is called a **Henkin constant**, and a sentence in H is called a **Henkin axiom**. Finally, $T \cup H$ is called the **Henkin extension** or the **Henkinization** of T.

Remark

The cardinality of C constructed in the above proof coincides with the larger one of the cardinality of $\mathcal{L}$ and the countable infinity.
In particular, if $\mathcal{L}$ is finite or countably infinite, C is countably infinite.

This follows from the following facts.

- Each C_{i+1} is no larger than the number of formulas in $\mathcal{L} \cup C_i$.
- The set of finite sequences from countably many symbols is countably infinite. The set of finite sequences from uncountable κ symbols has cardinality κ.

Lemma 3.7

Let T be a consistent theory in a language $\mathcal{L}$. Then, there is a theory S in $\mathcal{L}' \supset \mathcal{L}$, which satisfies the following conditions.

(0) $T \subseteq S$ and S is a consistent theory of $\mathcal{L}'$.

(1) For each sentence $\exists x \varphi(x)$ in $\mathcal{L}'$ such that $S \vdash \exists x \varphi(x)$, there exists a constant c of $\mathcal{L}'$ such that $S \vdash \varphi(\mathsf{c})$.

(2) For any sentence φ in $\mathcal{L}'$, $\varphi \in S$ or $\neg\varphi \in S$.

Such an S is called a **complete Henkin extension** of T.

Proof. By Zorn's lemma, S exists as a **maximal consistent set** that includes $T \cup H$. Here, we also show the construction of S by transfinite recursion.

- Let $\{\sigma_i\}_{i<\alpha}$ enumerate the sentences in the language $\mathcal{L}'$ with ordinal numbers $i < \alpha$. Note that α is not smaller than either the cardinality of $\mathcal{L}'$ or $\aleph_0$.
- Define an increasing sequence $\{S_i\}_{i<\alpha}$ of consistent sets as follows.

$$S_0 = T \cup H, \text{ where } H \text{ is the set of Henkin axioms,}$$
$$S_{i+1} = \begin{cases} S_i \cup \{\sigma_i\}, & \text{if } S_i \cup \{\sigma_i\} \text{ is consistent,} \\ S_i \cup \{\neg\sigma_i\}, & \text{otherwise,} \end{cases}$$
$$S_\beta = \bigcup_{i<\beta} S_i, \quad \text{if } \beta(<\alpha) \text{ is a limit ordinal.}$$

- Then, we let $S = \bigcup_{i<\alpha} S_i$.
- It is clear from the construction that S satisfies conditions (0) and (2). Since $S = S \cup H$, S also satisfies condition (1). □

Note that since S satisfies (2), for any sentence σ in $\mathcal{L}'$, we have

$$\sigma \in S \quad \Leftrightarrow \quad S \vdash \sigma$$

Theorem 3.8

Any consistent theory has a model.

Proof.

- For a consistent theory T, we will construct a model of T using the set of Henkin constants C and the complete Henkin extension S.
- First, define the equivalence relation $\approx$ on C as follows:

$$\mathtt{c} \approx \mathtt{d} \quad \Leftrightarrow \quad (\mathtt{c} = \mathtt{d}) \in S.$$

Obviously, $\approx$ is a congruence relation because S contains the axioms of equations.

- Let A be the set of equivalence classes $[\mathtt{c}] = \{\mathtt{d} \in C : \mathtt{c} \approx \mathtt{d}\}$.
- We define $\mathcal{L}$-structure $\mathfrak{A} = (A, \mathtt{f}^{\mathfrak{A}}, \ldots, \mathrm{R}^{\mathfrak{A}}, \ldots)$ by defining the interpretation of the function symbol $\mathtt{f}$ and the relational symbol R of $\mathcal{L}$ as follows:

$$\mathtt{f}^{\mathfrak{A}}([\mathtt{c}_0], [\mathtt{c}_1], \ldots, [\mathtt{c}_{m-1}]) = [\mathtt{d}] \quad \Leftrightarrow \quad (\mathtt{f}(\mathtt{c}_0, \mathtt{c}_1, \ldots, \mathtt{c}_{m-1}) = \mathtt{d}) \in S,$$

$$\mathrm{R}^{\mathfrak{A}}([\mathtt{c}_0], [\mathtt{c}_1], \ldots, [\mathtt{c}_{n-1}]) \quad \Leftrightarrow \quad \mathrm{R}(\mathtt{c}_0, \mathtt{c}_1, \ldots, \mathtt{c}_{n-1}) \in S.$$

The well-definedness can be derived from the fact that $\approx$ is a congruence relation.

Proof (continued).

Now, we prove the following by induction on the construction of formula.

- For any formula $\varphi(x_0, x_1, \ldots, x_{n-1})$ in $\mathcal{L}$ (with all free variables displayed).

$$\varphi([\mathsf{c}_0], [\mathsf{c}_1], \ldots, [\mathsf{c}_{n-1}]) \in \mathrm{Th}(\mathfrak{A}_A) \quad \Leftrightarrow \quad \varphi(\mathsf{c}_0, \mathsf{c}_1, \ldots, \mathsf{c}_{n-1}) \in S$$

(1) If φ is an atomic formula, it immediately follows from the definition of $\mathfrak{A}$.
(2) If $\varphi \equiv \psi \vee \theta$,

$$\begin{aligned} \psi \vee \theta \in \mathrm{Th}(\mathfrak{A}_A) &\Leftrightarrow \psi \in \mathrm{Th}(\mathfrak{A}_A) \text{ or } \theta \in \mathrm{Th}(\mathfrak{A}_A) \\ &\Leftrightarrow \psi \in S \text{ or } \theta \in S \quad \Leftrightarrow \quad \psi \vee \theta \in S. \end{aligned}$$

(3) If $\varphi \equiv \exists x \psi(x)$,

$$\begin{aligned} \exists x \psi(x) \in \mathrm{Th}(\mathfrak{A}_A) &\Leftrightarrow \text{there exists } \mathsf{c} \text{ s.t. } \psi([\mathsf{c}]) \in \mathrm{Th}(\mathfrak{A}_A) \\ &\Leftrightarrow \text{there exists } \mathsf{c} \text{ s.t. } \psi(\mathsf{c}) \in S \\ &\Leftrightarrow \exists x \psi(x) \in S \ (\because \text{Condition (1) of } S). \end{aligned}$$

Other forms of formulas can be treated in the same way.

- Therefore, $\mathfrak{A}$ is a model of S, and it is also a model of T. □

Theorem 3.9 (Gödel's completeness theorem [1])

$T \vdash \varphi \Leftrightarrow T \models \varphi.$

Proof.

($\Rightarrow$). Suppose there exists a proof tree P of $T \vdash \varphi$.

- Let $\mathfrak{A}$ be a model of T. Treating the truth value of the sequent $\varphi_1, \ldots, \varphi_n$ as the truth value of $\varphi_1 \vee \cdots \vee \varphi_n$. A sequent at a top node of P is among the axioms of T, the law of excluded middle, and the axioms of equations, and thus is true in $\mathfrak{A}$.
- For each inference rule appearing in P, if the premises are true, the conclusion is also true. So, all sequents that appear in P are true. In particular, φ at the root is true.

($\Leftarrow$). Assume $T \vdash \varphi$ does not hold.

- If φ includes free variables $x_0, \ldots, x_n$, we treat it as $\forall x_0 \cdots \forall x_n \varphi$ for a sentence φ.
- By the lemma in Page 6, $T \cup \{\neg\varphi\}$ is consistent. Therefore, by the lemma in Page 14, $T \cup \{\neg\varphi\}$ has a model. Thus $T \models \varphi$ does not hold, which proves the completeness theorem.

[1] Exactly, Gödel-Henkin's completeness theorem. Gödel's original version was a little weaker.

Theorem 3.10 (Compactness theorem)

A theory T has a model if and only if any finite subset of T has a model.

Proof. $\Rightarrow$ is obvious and we only show $\Leftarrow$.

- By way of contradiction, suppose T has no model. Then, $T \models$ (the empty sequent) holds. By the completeness theorem, we also have $T \vdash$. Since a proof tree includes only finitely many axioms, there is a finite set $T' \subset T$ such that $T' \vdash$. Therefore, by the completeness theorem, $T' \models$, that is, some finite subset of T has no model. □

Problem 2

Prove that if $T_1 \cup T_2$ is inconsistent, then there exists a sentence σ such that $T_1 \vdash \sigma$ and $T_2 \vdash \neg\sigma$.

Problem 3

Show that there is no first-order theory that has arbitrarily large finite models and has no infinite models.

Theorem 3.11 (Löwenheim-Skolem's downward theorem)

A consistent theory in the language $\mathcal{L}$ has a model whose cardinality is less than or equal to the cardinality of $\mathcal{L}$ or the countable infinity.

Proof.

- By the proof of Completeness Theorem and the remark on Page 11, the cardinality of the set C of Henkin constants coincides with the larger one of the cardinality of $\mathcal{L}$ and the countable infinity.
- Since $\mathfrak{A}$ is constructed by taking an equivalence class from this, the cardinality of $|\mathfrak{A}|$ is less than or equal to the cardinality of $\mathcal{L}$ or the countable infinity.

□

Remarks.

- By the theorem, real number theory and set theory, which are axiomatized in countable languages, have countable models. It seems unreasonable at first glance that they have countable models since they deals with uncountable sets (**Skolem's paradox**).
- However, "A is countable" means that there is a one-to-one correspondence between A and the natural numbers. Therefore, if such a function exists outside the model, A would be uncountable in the model but it is countable outside.
- Historically, Skolem's discovery preceded Gödel's completeness theorem and revealed the relativity of the cardinality.

Theorem 3.12 (Löwenheim-Skolem-Tarski's upward theorem)

If a theory in a language $\mathcal{L}$ has an infinite model, then it has a model with arbitrary cardinality greater than or equal to the cardinality of $\mathcal{L}$.

Proof.

- Let $\mathfrak{A}$ be an infinite structure of the theory T_0 of $\mathcal{L}$ and κ a cardinal number greater than or equal to the cardinality of $\mathcal{L}$. Let C be a set of new constants with size κ.
- Let

$$T = T_0 \cup \{\mathtt{c} \neq \mathtt{d} : \mathtt{c} \text{ and } \mathtt{d} \text{ are two different constants belonging to } C\}$$

- Then, any finite subset T' of T has a model $\mathfrak{A}'$, which is constructed from $\mathfrak{A}$ with an appropriate interpretation of constants in C so that a finite number of $\mathtt{c} \neq \mathtt{d}$ contained in T' are true.
- Therefore, by the compactness theorem, T has a model. However, due to the properties of T, the cardinality of any model is greater than or equal to κ.
- On the other hand, by the downward theorem, since T has a model with cardinality $\leq \kappa$, it follows that there exists a model with exactly cardinality κ.

□

Existence of non-standard models of arithmetic

- Let $\mathcal{N} = (\mathbb{N}, 0, 1, +, \cdot, <)$ be the standard model of arithmetic (natural number theory).
- Let Th($\mathcal{N}$) := $\{\sigma : \mathcal{N} \models \sigma\}$. $\mathcal{N}$ is naturally a model of Th($\mathcal{N}$), but there also exist models of Th($\mathcal{N}$) that are not isomorphic to $\mathcal{N}$, which are called **nonstandard models** of arithmetic.
- Using the compactness theorem, we construct a nonstandard model of arithmetic as follows. First, with c as a new constant, for each $k \in \mathbb{N}$

$$T_k = \mathrm{Th}(\mathcal{N}) \cup \{0 < \mathrm{c}, 1 < \mathrm{c}, 1+1 < \mathrm{c}, 1+1+1 < \mathrm{c}, \dots, \overbrace{1+1+\cdots+1}^{k \text{ times}} < \mathrm{c}\}$$

- The structure of $\mathcal{N}$ plus the interpretation of the constant c as $k+1$ is a model of T_k.
- Let $T = \bigcup_{k \in \omega} T_k$. Any finite subset of T is contained in some T_k and so satisfiable. Hence, by the compactness theorem, T also has a model $\mathcal{M}$, where the value of c is larger than any standard natural number.
- That is, $\mathcal{M}$ has elements that are not standard natural numbers.
- By removing the constant c from the structure, $\mathcal{M}$ can be regarded as a non-standard model of arithmetic in the language $\mathcal{L}_{\mathrm{OR}}$.

§4. Conservative extension

Definition 4.1

Let T and T' be theories in languages $\mathcal{L}$ and $\mathcal{L}'$, respectively, and $\mathcal{L} \subset \mathcal{L}'$. Then, T' is called a **conservative extension** of T if for any sentence σ in $\mathcal{L}$, $T \vdash \sigma \iff T' \vdash \sigma$.

Example 6

Let T be a group theory in $\mathcal{L} = (\cdot)$ consisting of

$$\begin{cases} (x \cdot y) \cdot z = x \cdot (y \cdot z), \\ \exists z(\forall x(z \cdot x = x) \wedge \forall x \exists y(y \cdot x = z)). \end{cases}$$

Let T' be another group theory in $\mathcal{L}' = (\cdot, \mathsf{e}, {}^{-1})$ consisting of

$$\begin{cases} (x \bullet y) \bullet z & = & x \bullet (y \bullet z) & \text{(associativity)} \\ \mathsf{e} \bullet x & = & x & \text{(left identity)} \\ x^{-1} \bullet x & = & \mathsf{e} & \text{(left inverse)} \end{cases}$$

Then T' is a conservative extension of T (will be proved by the following theorem and corollary.).

Theorem 4.2

Let $\varphi(x_1, \ldots, x_n, y)$ be a formula in a language $\mathcal{L}$ with no free variables other than $x_1, \ldots, x_n, y$. If $T \vdash \forall x_1 \cdots \forall x_n \exists y \varphi(x_1, \ldots, x_n, y)$, then a theory extended with a new n-ary function symbol $\mathtt{f}$, $T' = T \cup \{\forall x_1 \cdots \forall x_n \varphi(x_1, \ldots, x_n, \mathtt{f}(x_1, \ldots, x_n))\}$ is a conservative extension of T.

Proof

- Let $\mathfrak{A}$ be any model of T. Suppose $T \vdash \forall x_1 \cdots \forall x_n \exists y\ \varphi(x_1, \ldots, x_n, y)$.
- Since $\mathfrak{A} \models \forall x_1 \cdots \forall x_n \exists y\ \varphi(x_1, \ldots, x_n, y)$, by axiom of choice, we construct a function $\mathtt{f}^{\mathfrak{A}}$ on $\mathfrak{A}$ such that

$$\mathfrak{A} \models \forall x_1 \cdots \forall x_n\ \varphi(x_1, \ldots, x_n, \mathtt{f}(x_1, \ldots, x_n))$$

- Let $\mathfrak{A}^*$ be a structure $\mathfrak{A} \cup \{\mathtt{f}^{\mathfrak{A}}\}$. Then it is clearly a model of T'.
- Now, take any theorem σ of T' in the language $\mathcal{L}$. Then it is true in $\mathfrak{A}^*$. Since σ does not include $\mathtt{f}$, its truth value is irrelevant to $\mathtt{f}^{\mathfrak{A}}$. Therefore, σ should also hold in $\mathfrak{A}$.
- Since $\mathfrak{A}$ is an arbitrary model of T, by the completeness theorem we have $T \vdash \sigma$. □

As a special case of this theorem, we have the following corollary.

Corollary 4.3

If $T \vdash \exists x \varphi(x)$, then $T \cup \{\varphi(c)\}$ with a new constant c is a conservative extension of T.

For a quantifier-free formula φ, $\forall x_1 \cdots \forall x_n \varphi$ is called a **$\forall$-formula** or **Π_1-formula**.

Theorem 4.4

Every theory T has a conservative extension theory T' consisting only of $\forall$-sentences.

Proof.
Let $\mathcal{L}$ be any language. For each formula of the form $\exists y \varphi(x_1, \ldots, x_n, y)$ in $\mathcal{L}$ with no free variables other than $x_1, \ldots, x_n$, we add a new function symbol $\mathtt{f}_{\exists y \varphi(x_1,\ldots,x_n,y)}$ and collect them as F_1. Let

$$S_1 = \{\forall x_1 \cdots \forall x_n (\exists y \varphi(x_1, \ldots, x_n, y) \leftrightarrow \varphi(x_1, \ldots, x_n, \mathtt{f}_{\exists y \varphi(x_1,\ldots,x_n,y)}(x_1, \ldots, x_n))) :$$

$$\exists y \varphi(x_1, \ldots, x_n, y) \text{ is a formula in } \mathcal{L}\}$$

By Theorem 4.2 in Page 22, for any theory T of $\mathcal{L}$, $T \cup S_1$ is a conservative extension of T.

Proof (continued).

- Next, for each formula of the form $\exists y \varphi(x_1, \ldots, x_n, y)$ in the language $\mathcal{L} \cup F_1$, we add a new function symbol and collect them as F_2 and similarly define S_2.

- By repeating this process, we finally put

$$F = \bigcup_{i \in \mathbb{N}} F_i, \quad S = \bigcup_{i \in \mathbb{N}} S_i$$

- Then, for any theory T in $\mathcal{L}$, $T \cup S$ is a conservative extension of T, which is called an (**iterated**) **Skolem extension** of T. A symbol belonging to F is called a **Skolem function**, and a sentence belonging to S is called a **Skolem axiom**.
  These are generalizations of Henkin constants and Henkin axioms in the proof of the completeness theorem.

- Under the Skolem axioms S, any formula φ in $\mathcal{L}' = \mathcal{L} \cup F$ is equivalent to a $\forall$-formula, which can be shown by induction on the construction of φ.

Proof (continued).

- Moreover, in order to prove that any formula is equivalent to a $\forall$-formula, we may restrict the Skolem axioms to the following set.

$$S' = \{\forall x_1 \ldots \forall x_n \forall y(\varphi(x_1, \ldots, x_n, y) \to \varphi(x_1, \ldots, x_n, \mathtt{f}_{\exists y \varphi(x_1,\ldots,x_n,y)}(x_1, \ldots, x_n))) :$$

$$\varphi(x_1, \ldots, x_n, y) \text{ is a quantifier-free formula of } \mathcal{L}'\}$$

Note here that all formulas in S' are $\forall$-sentences.

- Let us consider an example. First of all, we transform a formula into prenex normal form by pushing an inner quantifier forward. For instance, change $\theta \land \forall x \xi(x)$ to $\forall z(\theta \land \xi(z))$ by replacing the bound variable x with a new variable z if necessary.

- Now consider a formula $\exists x \forall y \exists z \theta(x, y, z)$ or $\exists x \neg \exists y \neg \exists z \theta(x, y, z)$. First replace z in $\theta(x, y, z)$ with $\mathtt{f}_{\exists z \theta(x,y,z)}(x, y) \in F_1$ and put the following into S_1

$$\forall x, y, z(\theta(x, y, z) \to \theta(x, y, \mathtt{f}_{\exists z \theta(x,y,z)}(x, y)).$$

- For simplicity, we write $\theta_1(x, y)$ for $\theta(x, y, \mathtt{f}_{\exists z \theta(x,y,z)}(x, y))$. Next, replace y in $\neg\theta_1(x, y)$ with $\mathtt{f}_{\exists y \neg \theta_1(x,y)}(x) \in F_2$ and put the following into S_2

$$\forall x, y(\neg\theta_1(x, y) \to \neg\theta_1(x, \mathtt{f}_{\exists y \neg \theta_1(x,y)}(x)).$$

- Again for simplicity, we write $\theta_2(x)$ for $\neg\theta_1(x, \mathtt{f}_{\exists y \neg\theta_1(x,y)}(x))$. Replace x in $\neg\theta_2(x)$ with a constant $\mathtt{f}_{\exists x \neg\theta_2(x)} \in F_3$ and put the following into S_3

$$\forall x(\neg\theta_2(x) \to \neg\theta_2(\mathtt{f}_{\exists x \neg\theta_2(x)})).$$

- Then under the assumption S_3, we have

$$\begin{aligned} \exists x \forall y \exists z \theta(x,y,z) &\leftrightarrow \exists x \forall y \theta_1(x,y) \\ &\leftrightarrow \exists x \neg\exists y \neg\theta_1(x,y) \\ &\leftrightarrow \exists x \neg\theta_2(x) \\ &\leftrightarrow \neg\theta_2(\mathtt{f}_{\exists x \neg\theta_2(x)}). \end{aligned}$$

  Thus, $\exists x \forall y \exists z \theta(x,y,z)$ is equivalent to a quantifier-free sentence.

- For each axiom (sentence) in the theory T, we rewrite it as a quantifier-free sentence in $\mathcal{L} \cup F$ and collect all of them as T''.

- Then $T' = T'' \cup S'$ is a conservative extension of T consisting of only $\forall$-sentences. □

Next, we consider an interpretation of a theory into a distinct language.

- First, we discuss a function symbol introduced by definition. As we will see later, this is a special case of interpretation.
- In the following, $\exists ! y \psi(y)$ is an abbreviation of

$$\exists y \psi(y) \wedge \forall y_1 \forall y_2 (\psi(y_1) \wedge \psi(y_2) \rightarrow y_1 = y_2),$$

  which means "there exists a unique y that satisfies $\psi(y)$."
- Let T be a theory in $\mathcal{L}$, and $\varphi(x_1, \ldots, x_n, y)$ a formula in $\mathcal{L}$ with no free variables other than $x_1, \ldots, x_n, y$. Assume $T \vdash \forall x_1 \cdots \forall x_n \exists ! y \varphi(x_1, \ldots, x_n, y)$. Then the theory $T' = T \cup \{\forall x_1 \cdots \forall x_n \forall y (\varphi(x_1, \ldots, x_n, y) \leftrightarrow \mathtt{f}(x_1, \ldots, x_n) = y)\}$ is called an **expansion** of T **by definition** .
- By Theorem 4.2 in Page 22, T' is a conservative extension of T.

Suppose $T' = T \cup \{\forall x_1 \cdots \forall x_n \forall y(\varphi(x_1, \ldots, x_n, y) \leftrightarrow \mathtt{f}(x_1, \ldots, x_n) = y)\}$. Given a formula ψ of $\mathcal{L} \cup \{\mathtt{f}\}$, we construct $\psi^{-\mathtt{f}}$ in $\mathcal{L}$ by the following procedure.

(1) If ψ does not include $\mathtt{f}$, then terminate this process by setting $\psi^{-\mathtt{f}} = \psi$.

(2) If ψ contains $\mathtt{f}$, take an atomic subformula θ containing it, and choose a subterm $\mathtt{f}(t_0, \ldots, t_{n-1})$ in it such that no t_i contains $\mathtt{f}$.

(3) In θ, replace the subterm selected in (2) with a new variable y and call it $\theta_1(y)$.

(4) Replace θ in ψ by $\exists y(\varphi(t_0, \ldots, t_{n-1}, y) \land \theta_1(y))$, and then we regard it as a new ψ, and then go to (1).

It is easy to see that $\psi^{-\mathtt{f}}$ thus constructed satisfies the following:

$$T' \vdash \psi \leftrightarrow \psi^{-\mathtt{f}}$$

Lemma 4.5

Let T be a theory in $\mathcal{L}$, φ a formula in $\mathcal{L}$, and θ be a subformula of φ. Assume $T \vdash \theta \leftrightarrow \theta'$. Let φ' be a formula obtained from φ by replacing some occurrence of θ in φ with θ'. Then $T \vdash \varphi \leftrightarrow \varphi'$.

Proof. By the completeness theorem, it is enough to show that in any model $\mathfrak{A}$ of T, if $\theta \leftrightarrow \theta'$ holds, then $\varphi \leftrightarrow \varphi'$ holds, which is straightforward from Tarski's truth definition clauses. □

Remark

- The same argument holds for relational symbols. Using a new relational symbol R,
$$T' = T \cup \{\forall x_1 \cdots \forall x_n (\varphi(x_1, \ldots, x_n) \leftrightarrow \mathrm{R}(x_1, \ldots, x_n))\},$$
is a conservative extension of T.
- Let $\psi^{-\mathrm{R}}$ denote the formula obtained from ψ by replacing all $\mathrm{R}(t_1, \ldots, t_n)$ with $\varphi(t_1, \ldots, t_n)$. Then
$$T' \vdash \psi \leftrightarrow \psi^{-\mathrm{R}}$$

Challenging problem

Let Σ be a theory in a language $\mathcal{L}$ including an n-ary relation symbol R and some others. Then, R is said to be explicitly definable in Σ, if there exists a formula $\varphi(x_0, \ldots, x_{n-1})$ in $\mathcal{L} - \{R\}$ such that

$$\Sigma \vdash \forall x_0, \ldots, x_{n-1}(R(x_0, \ldots, x_{n-1}) \leftrightarrow \varphi(x_0, \ldots, x_{n-1})).$$

Now, we construct Σ' from Σ by replacing all occurrences of R by a new symbol R'. Then, R is said to be implicitly definable in Σ, if the following holds

$$\Sigma \cup \Sigma' \vdash \forall x_0, \ldots, x_{n-1}(R(x_0, \ldots, x_{n-1}) \leftrightarrow R'(x_0, \ldots, x_{n-1})).$$

Show that R is explicitly definable in Σ iff R is implicitly definable in Σ.

Now we are ready to define a language interpretation.

Definition 4.6

Given two languages $\mathcal{L}, \mathcal{L}'$ and a theory T' of the language $\mathcal{L}'$. A pair $\langle U, I \rangle$ that satisfies the following conditions is called an **interpretation (translation)** (within T') of language $\mathcal{L}$.

(1) U is a one-variable formula in $\mathcal{L}'$. (It represents the domain of the theory in $\mathcal{L}$.)

(2) I is a function from formulas in $\mathcal{L}$ to formulas in $\mathcal{L}'$, and if $\mathtt{f}$ is an n-ary function symbol, $I(\mathtt{f})$ is an $(n+1)$-ary formula; if R is an n-ary relationship symbol, $I(\mathrm{R})$ is also an n-ary formula.

(3) $T' \vdash \exists x U(x)$.

(4) For all functional symbol f,

$$T' \vdash \forall x_1 \cdots \forall x_n (U(x_1) \land \cdots \land U(x_n) \to \exists! y (I(\mathtt{f})(x_1, \ldots, x_n, y) \land U(y))).$$

- Next, we will translate the $\mathcal{L}$-formulas. Since it is difficult to directly translate them into $\mathcal{L}'$-formulas, we extend $\mathcal{L}'$ by definition and consider the interpretation there.
- Let $\mathtt{f}$ be the function symbol defined by $I(\mathtt{f})$ with some modification, R the relational symbol defined by $I(\mathrm{R})$. The remaining is that $T \vdash \forall x_1 \cdots \forall x_n \exists! y I(\mathtt{f})(x_1, \ldots, x_n, y)$ does not necessarily hold. However, if we take an arbitrary constant a and set it as follows, $I'(\mathtt{f})$ satisfies the above condition. Then define the function $\mathtt{f}$ by this.

$$I'(\mathtt{f})(x_1, \ldots, x_n, y) \Leftrightarrow$$

$$((U(x_1) \wedge \cdots \wedge U(x_n)) \wedge I(\mathtt{f})(x_1, \ldots, x_n, y)) \vee ((\neg U(x_1) \vee \cdots \vee \neg U(x_n)) \wedge y = a).$$

- If we expand $\mathcal{L}'$ in this way, the terms of $\mathcal{L}$ will remain unchanged after interpretation, and so is the atomic formulas and the propositional connective symbols.
- The last thing to deal with the quantification symbols. If we denote the interpreted formula φ in $\mathcal{L}$ as φ^I,
  (1) $(\exists x \psi)^I$ is $\exists x (U(x) \wedge \psi^I)$.
  (2) $(\forall x \psi)^I$ is $\forall x (U(x) \to \psi^I)$.

Definition 4.7

- Let T and T' be theories of languages $\mathcal{L}$ and $\mathcal{L}'$, respectively, and $\langle U, I\rangle$ be an interpretation of language $\mathcal{L}$ in T'.
- Then for any sentence σ in $\mathcal{L}$, if
$$T \vdash \sigma \quad \Rightarrow \quad T' \vdash \sigma^I$$
then $\langle U, I\rangle$ is said to be an **interpretation** of theory T within T'.
- If such $\langle U, I\rangle$ exists, T is said to be **interpretable** within T'.
- Moreover, if the following holds
$$T \vdash \sigma \quad \Leftrightarrow \quad T' \vdash \sigma^I$$
$\langle U, I\rangle$ is called a **faithful interpretation** of T' in T.

Example 7

If T is an expansion of T', then there is a faithful interpretation $\langle U, I\rangle$.
Let $U(x)$ be $x = x$. For a defined function $\mathtt{f}$ or relation R, let $I(\mathtt{f})$ or $I(\mathrm{R})$ preserve its definition. And so are the interpretations of other symbols.

Example 8

Let $\mathfrak{N} = (\mathbb{N}, +, \cdot, 0, 1, <)$, $\mathfrak{Z} = (\mathbb{Z}, +, \cdot, 0, 1)$.
There exists a faithful interpretation $\langle U, I\rangle$ from Th($\mathfrak{N}$) to Th($\mathfrak{Z}$):

$$U(x) \equiv \exists x_1 \exists x_2 \exists x_3 \exists x_4 (x = x_1 \cdot x_1 + \cdots + x_4 \cdot x_4)$$

$$I(+)(l, m, n) \equiv l + m = n, \quad I(\cdot)(l, m, n) \equiv l \cdot m = n$$

$$I(0)(n) \equiv n = 0, \quad I(1)(n) \equiv n = 1$$

$$I(<)(m, n) \equiv \exists x U(x) \wedge x \neq 0 \wedge m + x = n$$

Problem 4

Show that there exists a translation $\langle U, I \rangle$ from $\mathrm{Th}(\mathfrak{Z})$ that is faithful to $\mathrm{Th}(\mathfrak{N})$.

Problem 5

1. Show that Peano arithmetic PA is interpretable within ZF set theory.
2. Show that ZF without Infinity axiom is interpretable within PA.

Homework # 2

Solve Problem 4 or 5.

Remark.

- If a faithful translation from T to T' exists, provability in T is reduced to that of T'.
- Therefore, if the provability of T' is attainable (for example, if T' is decidable), then so is T.
- Another powerful argument is to show the undecidability of T' by translating the theory T.

Thank you for your attention!