INTRODUCTION TO FIRST ORDER LOGIC AND MODEL THEORY

BILL LAI

Abstract. This paper is a detailed introduction to the fundamental concepts and results in first order logic and model theory for readers who have some background in propositional logic and probably first order logic. Besides providing the definitions and proving the theorems, I will also try to uncover the motivations, rationale, and implications behind them to promote a deeper understanding in logic and formal language.

Contents

1. Introduction: First Order Language

A serious inquiry of mathematical logic appears challenging even at the beginning. Logic drives inferences from one mathematical statement to another; therefore, the study of logic is the study of the relations of these statements. This means that, however, any direct examination would be difficult. Consider the statements "if $x = 1$, then $y = 1$" and "the valuation of y at 1 is necessary for x to be 1". For many mathematicians, they are indeed logically equivalent, but even if so, such equivalence is hard to be depicted formally and rigorously. Moreover, for some people, even the equivalence is subject to doubt, as it could be argued that what "necessary" means is open to interpretation.

It is not difficult to see that the problem above arises from the nature of our ordinary language that is used to represent the mathematical statements. Complicated sentential structures and indefinite interpretations of vocabulary could conceal statements' intension and logical nature. As a result, the relations between mathematical statements are not always clear.

A natural solution is to develop a new language that can both 1) represent mathematical statements in a more systematic and simplified way and 2) uncover their logical nature. As a result, the **first order language** emerges as one of the most desirable candidates, and this paper will partially illustrate the reasons.

I assume the readers have the knowledge of the lexicon and syntax of the first order language, so I will not define the language formally. However, I want to make the following remarks about the notations used in this paper at this point.

Remarks 1.1. (1) I use $\mathcal{L}$ to denote the first order language in general, and $\mathcal{L}_{\mathcal{A}}$ to denote the specific first order language whose non-logical symbols (i.e., constants, functions, and relations) constitute the set $\mathcal{A}$.

(2) I use c as the symbol for constant, F for function, P for relation, x for variable, and $\equiv$ for equivalence.

(3) I use $\tau(x_1, \dots, x_n)$ as the symbol for the term τ, where every variable occurring in τ is in $\{x_1, \dots, x_n\}$, and $\varphi(x_1, \dots, x_n)$ for the formula φ, where every variable occurring freely in φ is in $\{x_1, \dots, x_n\}$. I use $\tau(x_1, \dots, x_n; \tau_1, \dots, \tau_n)$ for the result of replacing every occurrence of x_i in τ with term τ_i, and $\varphi(x_1, \dots, x_n; \tau_1, \dots, \tau_n)$ for the result of replacing every free occurrence of x_i in φ with term τ_i.

(4) For simplicity, I use "$\mathcal{L}$-" to describe objects of the language $\mathcal{L}$. For instance, an "$\mathcal{L}$-term" is a term of the language $\mathcal{L}$, and an "$\mathcal{L}$-formula" is a formula of the language $\mathcal{L}$.

The following theorems, which show that $\mathcal{L}$ is unambiguous, are conceptually crucial as unambiguity is necessary for any language to be capable of representing mathematical statements and to be a promising tool for our inquiry of mathematical logic.

Theorem 1.2. *If τ is a term, then exactly one of the following holds:*

(1) τ is some unique variable x;

(2) τ is some unique constant c;

(3) τ is $F(\tau_1, \dots, \tau_n)$, where F is some unique n-variable function, and $\langle \tau_1, \dots, \tau_n \rangle$ is some unique sequence of terms.[1]

Theorem 1.3. *If φ is a formula, then exactly one of the following holds:*

(1) φ is $(\tau_1 \equiv \tau_2)$, where $\langle \tau_1, \tau_2 \rangle$ is some unique sequence of terms.

(2) φ is $P(\tau_1, \dots, \tau_n)$, where P is some unique n-variable relation, and $\langle \tau_1, \dots, \tau_n \rangle$ is some unique sequence of terms.

(3) φ is $(\neg\psi)$, where ψ is some unique formula.

(4) φ is $(\varphi_1 \vee \varphi_2)$, where $\langle \varphi_1, \varphi_2 \rangle$ is some unique sequence of formulas.

(5) φ is $(\varphi_1 \wedge \varphi_2)$, where $\langle \varphi_1, \varphi_2 \rangle$ is some unique sequence of formulas.

(6) φ is $(\varphi_1 \to \varphi_2)$, where $\langle \varphi_1, \varphi_2 \rangle$ is some unique sequence of formulas.

(7) φ is $(\forall x\psi)$, where ψ is some unique formula.

(8) φ is $(\exists x\psi)$, where ψ is some unique formula.[2]

In the following sections, I will introduce the basic concepts in first order model theory, including structure, deduction system, and theory. I will build up these concepts from the bottom so that the readers can understand the rationale and motivations behind. My introduction will show that these concepts and their related theorems are natural consequences of a serious inquiry of the first order mathematical logic. It will shed light on the strengths and limitations of logic and formal language, on what they can do and what they cannot do.

---

[1] For proof, see page 47 of [2].

[2] See page 50 of [2].

2. First Order Structure

It seems that $\mathcal{L}$, by itself, is still insufficient for our purpose. Having the means to decide the truth of a statement is crucial for studying logic, as logic aims to deduce true statements from true statements. However, considering the $\mathcal{L}$-formula $F(x_1, x_2) \equiv x_3$, which represents the statement $x_1 + x_2 = x_3$, we find that we have no means to determine its truth. Specifically, the original statement is true when $x_1 = 1$, $x_2 = 2$, and $x_3 = 3$, while false when $x_1 = 1$, $x_2 = 1$, and $x_3 = 0$, but this property is not captured by the $\mathcal{L}$-formula at all — we have not been able even to tell what $F(x_1, x_2) \equiv x_3$ is when $x_1 = 1$, $x_2 = 1$, and $x_3 = 0$, or when $x_1 = 1$, $x_2 = 1$, and $x_3 = 0$!

The problem above reveals a shortcoming of $\mathcal{L}$: it is formal without meaning, thereby unable to replicate the abundant connotations of mathematical statements. As a result, by $\mathcal{L}$ only it seems impossible to decide the truth of any statement. A natural solution to this problem is to assign to each non-logical symbol a specific meaning: if we somehow know that F is "addition", x_1 is 1, x_2 is 2, and x_3 is 3, for example, then we will be able to determine the truth of $F(x_1, x_2) \equiv x_3$.

However, beforehand we should also beware that the formality of $\mathcal{L}$ is exactly what allows it to uncover the underlying logic of mathematical statements. For instance, the commutativity of addition and multiplication are logically identical as they describe the same operational property, and they are both represented by $\mathcal{L}$ as $(\forall x_1(\forall x_2 F(x_1, x_2) \equiv F(x_2, x_1)))$. Therefore, we ought to keep $\mathcal{L}$ itself intact. In other words, the assignment of meanings cannot be done *within* the language; instead, it should be defined *externally* to $\mathcal{L}$. This leads to the following definitions:

Definition 2.1. A **structure** of a first order language $\mathcal{L}_\mathcal{A}$ is an ordered pair (M, I) satisfying the following:

(1) M, which is called the structure's **universe**, is a nonempty set;

(2) I, which is called the structure's **interpretation**, is a mapping defined on $\mathcal{A}$ such that:

(a) $I(c) \in M$ for all constants $c \in \mathcal{A}$;

(b) $I(F)$ is a function from M^n to M for all n-variable functions $F \in \mathcal{A}$;

(c) $I(P) \subseteq M^n$ for all n-variable relations $P \in \mathcal{A}$.

It would be beneficial to give some examples before introducing other definitions:

Example 2.2. Let $\mathcal{A} = \{c_0, F_+\}$, where c_0 is a constant and F_+ is a 2-variable function. Let $M = \mathbb{Z}, I(c_0) = 0$ and $I(F_+)$ be the addition function of integers. Then $\mathcal{M} = (M, I)$ is the structure of additive group of integers.

Example 2.3. Let $\mathcal{A} = \{P_<\}$, where $P_<$ is a 2-variable relation. Let $M = \mathbb{R}$ and $I(P_<)$ be the standard order on $\mathbb{R}$. Then $\mathcal{M} = (M, I)$ is the structure of linear order on $\mathbb{R}$.

These examples reveal how a structure "contextualizes" a language. Its universe specifies the objects that the language talks about, while its interpretation assigns specific meanings to the language's formal symbols. Thus, it allows us to say, for instance, that F means "addition" in the formula $F(x_1, x_2) \equiv x_3$. However, to be able to say that it is true or not, we still need one more step: to assign values to x_1, x_2, x_3. Thus, we need the following definition:

Definition 2.4. Let $\mathcal{M} = (M, I)$ be an $\mathcal{L}_\mathcal{A}$-structure. We call a function v an **$\mathcal{M}$-valuation** iff v is a mapping from the set of all variables to M.

In fact, this definition leads to a more general one, by expanding the domain from the set of variables to the set of terms.

Definition 2.5. Let $\mathcal{M} = (M, I)$ be an $\mathcal{L}_{\mathcal{A}}$-structure, and v be an $\mathcal{M}$-valuation. We define $\bar{v}$ as a mapping from the set of all $\mathcal{L}_{\mathcal{A}}$-terms to M such that:
(a) $\bar{v}(x) = v(x)$ for all variables x;
(b) $\bar{v}(c) = I(c)$ for all constants c;
(c) $\bar{v}(F(\tau_1, \ldots \tau_n)) = I(F)(\bar{v}(\tau_1), \ldots \bar{v}(\tau_n))$ for all n-variable functions F and terms $\tau_1, \ldots \tau_n$.

We can define a relation between two valuations:

Definition 2.6. Let $\mathcal{M}$ be an $\mathcal{L}_{\mathcal{A}}$-structure, τ be an $\mathcal{L}_{\mathcal{A}}$-term, and φ be an $\mathcal{L}_{\mathcal{A}}$-formula. We also let v and μ be two $\mathcal{M}$-valuations.
(1) v and μ are **consistent** on τ, denoted as $v \equiv_\tau \mu$, iff for any variable x that occurs in τ, $v(x) = \mu(x)$.
(2) v and μ are **consistent** on φ, denoted as $v \equiv_\varphi \mu$ or $v = \mu \bmod (\varphi)$, iff for any variable x that occurs freely in φ, then $v(x) = \mu(x)$.

Now we have fully developed the means to determine the truth of $F(x_1, x_2) \equiv x_3$ and all other $\mathcal{L}$-formulas. In sum, a structure creates a specific context for $\mathcal{L}$ that is necessary for us to decide the truth of statements, while a valuation determines the truth of a particular statement within a structure. The following definitions further reveal this fact.

Definition 2.7. Let $\mathcal{M} = (M, I)$ be an $\mathcal{L}_{\mathcal{A}}$-structure, v be an $\mathcal{M}$-valuation, and φ be an $\mathcal{L}_{\mathcal{A}}$-formula. We define that $\mathcal{M}$ **satisfies** φ by v, or simply $(\mathcal{M}, v)$ **satisfies** φ, denoted as

$$(\mathcal{M}, v) \vDash \varphi,$$

according to the complexity of φ:
(1) If φ is $(\tau \equiv \sigma)$, where τ and σ are two $\mathcal{L}_{\mathcal{A}}$-terms, then $(\mathcal{M}, v) \vDash \varphi$ iff $\bar{v}(\tau) = \bar{v}(\sigma)$.
(2) If φ is $P(\tau_1, \ldots, \tau_n)$, where $P \in \mathcal{A}$ is a n-variable relation, $\tau_1, \ldots, \tau_n$ are n $\mathcal{L}_{\mathcal{A}}$-terms, then $(\mathcal{M}, v) \vDash \varphi$ iff $\langle \bar{v}(\tau_1), \ldots, \bar{v}(\tau_n) \rangle \in I(P)$.
(3) If φ is $(\neg\psi)$, then $(\mathcal{M}, v) \vDash \varphi$ iff $(\mathcal{M}, v) \nvDash \psi$, where $(\mathcal{M}, v) \nvDash \psi$ means $(\mathcal{M}, v)$ does not satisfy ψ.
(4) If φ is $(\psi \wedge \theta)$, then $(\mathcal{M}, v) \vDash \varphi$ iff $(\mathcal{M}, v) \vDash \psi$ and $(\mathcal{M}, v) \vDash \theta$.
(5) If φ is $(\psi \vee \theta)$, then $(\mathcal{M}, v) \vDash \varphi$ iff $(\mathcal{M}, v) \vDash \psi$ or $(\mathcal{M}, v) \vDash \theta$.
(6) If φ is $(\psi \to \theta)$, then $(\mathcal{M}, v) \vDash \varphi$ iff $(\mathcal{M}, v) \nvDash \psi$ or $(\mathcal{M}, v) \vDash \theta$.
(7) If φ is $(\forall x\psi)$, then $(\mathcal{M}, v) \vDash \varphi$ iff for all $\mathcal{M}$-valuations μ, if $v \equiv_\varphi \mu$, then $(\mathcal{M}, \mu) \vDash \psi$.
(8) If φ is $(\exists x\psi)$, then $(\mathcal{M}, v) \vDash \varphi$ iff there exists an $\mathcal{M}$-valuation μ such that $v \equiv_\varphi \mu$ and $(\mathcal{M}, \mu) \vDash \psi$.

Before we proceed to Definition 2.8, which comes naturally from Definition 2.7, it might be beneficial to make a conceptual remark: structures, to be precise, are not mathematical contexts, but merely their representations, for the former, though external to $\mathcal{L}$, depend on it, while the latter do not. For instance, in Example 2.2, the structure defined exists only after we specify $\mathcal{A} = \{c_0, F_+\}$, while the context, i.e., the additive group of integers, always exists; therefore, the former is just the semiotic representation of the latter in $\mathcal{L}_{\mathcal{A}}$. However, for simplicity, I will still equate structures to "contexts" in most of the following discussion, and I will use "mathematical context" to refer to the latter at places where the distinction is important.

Definition 2.8. Let Γ be a set of $\mathcal{L}_{\mathcal{A}}$-formulas. Γ is **satisfiable** iff there exists a structure $\mathcal{M}$ and a valuation v within $\mathcal{M}$ such that

$$(\mathcal{M}, v) \vDash \varphi$$

for all formulas $\varphi \in \Gamma$.

Example 2.9. Let $\mathcal{M}$ be the structure defined in Example 2.2. Let $\varphi = (F_+(x, c_0) \equiv c_0), v_1(x) = 0, v_2(x) = 1$, then φ is satisfied in $\mathcal{M}$ by v_1, but not v_2.

Example 2.10. Let $\mathcal{M}$ be the structure defined in Example 2.3. Let $\varphi = P_<(x_1, x_2)$, $v_1(x_1) = v_2(x_1) = 0, v_1(x_2) = 1, v_2(x_2) = -1$, then φ is satisfied in $\mathcal{M}$ by v_1, but not v_2.

By defining satisfiability, we essentially define what it means to be *true* in $\mathcal{L}$: an $\mathcal{L}$-formula is "true" in the sense that it can be satisfied in some contexts. However, we should also note that the "truth" defined here is only one type of truth. In fact, in mathematics, there are multiple types of truth in terms of generality. Consider addition on $\mathbb{Z}$. $a + b = c$ is true for only certain values of a, b, and c, while $a + b = b + a$ holds for all $a, b \in \mathbb{Z}$; the "truth" in the latter case is more general than that in the former. Therefore, to fully capture the idea of "truth", the notion of satisfiability is not enough.

Definition 2.11. (1) For a given first order language $\mathcal{L}$, an $\mathcal{L}$-formula φ is **valid** in an $\mathcal{L}$-structure $\mathcal{M}$, denoted as $\mathcal{M} \vDash \varphi$, iff for all valuations v within $\mathcal{M}$,

$$(\mathcal{M}, v) \vDash \varphi.$$

In this case, we say $\mathcal{M}$ is a **model** of φ.

(2) For a given first order language $\mathcal{L}$, a set of $\mathcal{L}$-formulas Γ is **valid** in an $\mathcal{L}$-structure $\mathcal{M}$, denoted as $\mathcal{M} \vDash \Gamma$, iff for all valuations v within $\mathcal{M}$ and for all $\varphi \in \Gamma$,

$$(\mathcal{M}, v) \vDash \varphi.$$

In this case, we say $\mathcal{M}$ is a **model** of Γ.

Definition 2.12. (1) For a given first order language $\mathcal{L}$, an $\mathcal{L}$-formula φ is **universally valid**, denoted as $\vDash \varphi$, iff it is valid in every $\mathcal{L}$-structure.

(2) For a given first order language $\mathcal{L}$, a set of $\mathcal{L}$-formulas Γ is **universally valid**, denoted as $\vDash \Gamma$, iff it is valid in every $\mathcal{L}$-structure.

Note that universal validity implies validity, and validity implies satisfiablity, but the reverse are not true. However, there is one special type of formulas for which satisfiability is equivalent to validity. Because of such property, these formulas deserve special attention in our study.

Definition 2.13. An $\mathcal{L}$-formula φ is a **sentence** iff it does not have any free variables.

Theorem 2.14. *If a sentence φ is satisfied in structure $\mathcal{M}$ by some valuation v, then φ is valid in $\mathcal{M}$.*[3]

[3]For proof, see page 57 of [2].

There is another way to see that universal validity is stronger than satisfiablity and validity: while satisfiability and validity depend on the structure, universal validity does not. This perspective indicates a way to classify different types of truths among $\mathcal{L}$-formulas. Following this reasoning, we can think of another type of truth, which not only depends on the model, but also depends on other formulas.

Definition 2.15. For a set Γ of $\mathcal{L}$-formulas and $\varphi \in \Gamma$, φ is a **logical implication** of Γ, denoted as

$$\Gamma \vDash \varphi,$$

iff for all $\mathcal{L}$-structures $\mathcal{M}$ and for all $\mathcal{M}$-valuations v, if $(\mathcal{M}, v) \vDash \Gamma$, then $(\mathcal{M}, v) \vDash \varphi$.

The following remarks include some immediate consequences of the above definition:

Remarks 2.16. (1) A formula φ is a logical implication of a set of formulas Γ iff the satisfiablity of Γ ensures that of φ.

(2) A sentence ψ is a logical implication of a set of sentences Γ iff the validity of Γ ensures that of ψ; that is, any model of Γ is a model of ψ.

(3) A formula is universally valid iff it is a logical implication of the empty set.

(4) Any two universally valid formulas are logical implications of each other.

However, that universal validity does not depend on the structure seems to contradict the idea that structures enable the determination of truths. Such contradiction reminds us that $\mathcal{L}$ is not entirely formal. Although its non-logical symbols do not have specific meanings, the logical symbols (e.g. $\equiv, \cup, \neg \dots$) do. Therefore, when we previously said that $\mathcal{L}$ is meaningless, we were saying, precisely, that it does not have any *mathematical* meanings, but $\mathcal{L}$ itself does bear *logical* meanings. Thus, structures enable the determination of truths by giving $\mathcal{L}$-formulas mathematical meanings.

Therefore, a universally valid formula, by itself, is merely a logical, but not mathematical, truth. To make it a mathematical truth, we still need to place it under a specific context, i.e., structure. For instance, the universally valid formula $\forall x(x \equiv x)$ becomes a mathematical truth "for all x in $\mathbb{R}$, $x = x$" only after it is placed under the context $(\mathbb{R}, <)$. That is, its mathematical truth still depends on the structure. And since we are using $\mathcal{L}$ to study mathematical statements, we shall only focus on the mathematical truth of $\mathcal{L}$-formulas. As a result, despite the presence of universally valid formulas, we are still justified in asserting that only through structures and valuations can we determine the truth of $\mathcal{L}$-formulas. In fact, this assertion can also be grasped by the Replacement Theorem. To attain the theorem, we only need to introduce one more notation:

Definition 2.17. Let $\mathcal{M} = (M, I)$ be an $\mathcal{L}_{\mathcal{A}}$-structure, $b_1, \dots, b_n \in M$, $\tau(x_1, \dots, x_n)$ be an $\mathcal{L}_{\mathcal{A}}$-term, and $\phi(x_1, \dots, x_n)$ be an $\mathcal{L}_{\mathcal{A}}$-formula.

(1) $\tau[b_1, \dots, b_n]$ is defined to be $\bar{\mu}(\tau)$, where $\bar{\mu}$ is generated by the valuation μ and $\mu(x_i) = b_i (1 \leq i \leq n)$.

(2) $\mathcal{M} \vDash \varphi[b_1, \dots, b_n]$ iff for any $\mathcal{M}$-valuation μ that valuates variables $x_1, \dots, x_n$ at $b_1, \dots, b_n$ respectively, $(\mathcal{M}, \mu) \vDash \varphi$.

Now we may state the theorem.

Theorem 2.18. *(Replacement Theorem) Let $\mathcal{M} = (M, I)$ be a $\mathcal{L}_{\mathcal{A}}$-structure.*

(1) Let $\tau(x_1, \ldots, x_n), \tau_1, \ldots, \tau_n$ be $n+1$ $\mathcal{L}_{\mathcal{A}}$-terms, and $b_1, \ldots, b_n \in M$. If v is an $\mathcal{M}$-valuation such that $b_i = \bar{v}(\tau_i), 1 \leq i \leq n$, then

$$\bar{v}(\tau(x_1, \ldots, x_n; \tau_1, \ldots, \tau_n)) = \tau[b_1, \ldots, b_n].$$

(2) Let $\varphi(x_1, \ldots, x_n)$ be a $\mathcal{L}_{\mathcal{A}}$-formula, $\tau_1, \ldots, \tau_n$ be n $\mathcal{L}_{\mathcal{A}}$-terms, and $b_1, \ldots, b_n \in M$. If every τ_i can replace $x_i, 1 \leq i \leq n$, in

$$\varphi(x_1, \ldots, x_n)$$

and v is an $\mathcal{M}$-valuation such that $b_i = \bar{v}(\tau_i), 1 \leq i \leq n$, then

$$(\mathcal{M}, v) \vDash \varphi(x_1, \ldots, x_n; \tau_1, \ldots, \tau_n) \iff (\mathcal{M}, v) \vDash \varphi[b_1, \ldots, b_n].$$[4]

The Replacement Theorem shows that replacing terms by terms does not change the truth or, equivalently, meaning of a formula, and what affects its truth is *the valuation of terms within a specific structure*. This discovery echoes with what we have already discussed: $\mathcal{L}$ itself has no mathematical connotation, and structures make individual formulas mathematically meaningful. This idea implies that a further inquiry would require us to divert from individual formulas to the more essential, underlying structures that contextualize them. Such direction would require us to examine structures in general and study different structures simultaneously, for if we just focus on one particular structure, our scope would be limited, and our focus would ultimately fall back to individual formulas.

To study multiple structures, it is not only natural but also crucial to understand their relations. **Isomorphism** appears to be one of the most fundamental.

Definition 2.19. Let $\mathcal{M} = (M, I), \mathbb{N} = (N, J)$ be two $\mathcal{L}_{\mathcal{A}}$-structures.

(1) The bijection $e : M \to N$ is an **isomorphism** between M and N iff e satisfies the following:

(a) if $c \in \mathcal{A}$ is a constant, then $e(I(c)) = J(c)$;

(b) if $F \in \mathcal{A}$ is an n-variable function, then for all $(a_1, \ldots, a_{n+1}) \in M^{n+1}$,

$$I(F)(a_1, \ldots, a_) = a_{n+1} \iff J(F)(e(a_1), \ldots, e(a_n)) = e(a_{n+1});$$

(c) if $P \in \mathcal{A}$ is an n-variable relation, then for all $(a_1, \ldots, a_n) \in M^n$,

$$(a_1, \ldots, a_n) \in I(P) \iff (e(a_1), \ldots, e(a_n)) \in J(P).$$

(2) $\mathcal{M}$ and $\mathcal{N}$ are **isomorphic**, denoted as $\mathcal{M} \cong \mathcal{N}$, iff there exists an isomorphism between $\mathcal{M}$ and $\mathcal{N}$.

A consequence of isomorphism is illustrated by the following theorem.

Theorem 2.20. *(Homogeny Theorem) Let $e : M \to N$ be an isomorphism between $\mathcal{L}$-structures $\mathcal{M} = (M, I)$ and $\mathcal{N} = (N, J)$. Then for any $\mathcal{M}$-valuation v, the composition of e and v, $e \circ v$, is an $\mathcal{N}$-valuation, and for any $\mathcal{L}$-formula ϕ,*

$$(\mathcal{M}, v) \vDash \phi \iff (\mathcal{N}, e \circ v) \vDash \phi.$$

Proof. (1) If x is a variable, then $v(x) \in M$, so $e \circ v(x) \in N$. Therefore, $e \circ v$ is an $\mathcal{N}$-valuation.

(2) We prove by induction that e maintains the valuation of terms by $\bar{v}$ and $e \,\bar{\circ}\, v$, that is,

$$e(\bar{v}(\tau)) = \overline{e \circ v}(\tau)$$

[4] For proof, see page 60 of [2]

for any term τ. Indeed, this relation holds if τ is a variable. If τ is a constant c, $\bar{v}(c) = I(c), \overline{e \circ v}(c) = J(c)$. Since $e(I(c)) = J(c)$, $e(\bar{v}(c)) = \overline{e \circ v}(c)$.

If τ is $F(\tau_1, \ldots, \tau_n)$, where F is an n-variable function and $\tau_1, \ldots, \tau_n$ are terms. By induction hypothesis, the proposition holds for $\tau_1, \ldots, \tau_n$. Therefore,

$$\begin{aligned}
\overline{e \circ v}(\tau) &= \overline{e \circ v}(F(\tau_1, \ldots, \tau_n)) \\
&= J(F)(e \mathbin{\bar{\circ}} v(\tau_1), \ldots, e \mathbin{\bar{\circ}} v(\tau_n)) \\
&= J(F)(e(\bar{v}(\tau_1)), \ldots, e(\bar{v}(\tau_n))) \\
&= e(I(F)(\tau_1, \ldots, \tau_n) \\
&= e(\bar{v}(\tau)).
\end{aligned}$$

(3) Finally, we prove that e maintains the satisfiablity of any formula φ by induction, which is the theorem.

If φ is $(\tau \equiv \sigma)$, where τ and σ are terms. Then

$$(\mathcal{M}, v) \vDash \varphi \iff \bar{v}(\tau) = \bar{v}(\sigma),$$

$$(\mathcal{N}, e \circ v) \vDash \varphi \iff \overline{e \circ v}(\tau) = \overline{e \circ v}(\sigma).$$

Since e is a bijection,

$$\bar{v}(\tau) = \bar{v}(\sigma) \iff e(\bar{v}(\tau)) = e(\bar{v}(\sigma)).$$

By conclusion in (2),

$$e(\bar{v}(\tau)) = e(\bar{v}(\sigma)) \iff \overline{e \circ v}(\tau) = \overline{e \circ v}(\sigma).$$

Therefore,

$$(\mathcal{M}, v) \vDash (\tau \equiv \sigma) \iff (\mathcal{N}, e \circ v) \vDash (\tau \equiv \sigma).$$

If φ is $P(\tau_1, \ldots, \tau_n)$, where P is an n-variable relation and $\tau_1, \ldots, \tau_n$ are terms. Then by induction hypothesis,

$$(\mathcal{M}, v) \vDash \varphi \iff (\bar{v}(\tau_1), \ldots, \bar{v}(\tau_n)) \in I(P),$$

and

$$(\mathcal{N}, e \circ v) \vDash \varphi \iff (\overline{e \circ v}(\tau_1), \ldots, \overline{e \circ v}(\tau_n)) \in J(P).$$

By the similar reasoning in the previous case, we prove that the theorem holds for φ.

If φ is $(\neg\psi)$, where ψ is a formula for which

$$(\mathcal{M}, v) \vDash \psi \iff (\mathcal{N}, e \circ v) \vDash \psi.$$

By definition we have

$$(\mathcal{M}, v) \vDash (\neg\psi) \iff (\mathcal{M}, v) \nvDash \psi,$$

and

$$((\mathcal{N}, e \circ v) \vDash (\neg\psi) \iff ((\mathcal{N}, e \circ v) \nvDash \psi.$$

Therefore, the theorem holds for φ.

If φ is $\psi_1 \to \psi_2$, by the similar reasoning in the previous case, we may prove that the theorem holds for φ.

Lastly, if φ is $(\forall x \psi)$, then by definition,

$$(\mathcal{M}, v) \vDash \varphi$$

iff for any $\mathcal{M}$-valuation μ, if $v \equiv_\varphi \mu$, then $(\mathcal{M}, \mu) \vDash \psi$.

We first suppose $(\mathcal{M}, v) \vDash \varphi$. Let $\mu^* \equiv_\varphi e \circ v$ be an $\mathcal{N}$-valuation. Since e is a bijection, we define

$$\mu(x) = e^{-1}(\mu^*(x)).$$

Then μ is an $\mathcal{M}$-valuation and $\mu^* = e \circ v$, thereby $\mu \equiv_\varphi v$. Therefore, $(\mathcal{M}, v) \vDash \psi$, and $(\mathcal{N}, e \circ v) \vDash \psi$ by induction hypothesis.

We then suppose $(\mathcal{M}, v) \nvDash \varphi$. We let μ be a valuation such that $\mu \equiv_\varphi v$, but $(\mathcal{M}, \mu) \nvDash \psi$. Then $e \circ \mu \equiv_\varphi e \circ v$. By induction hypothesis,

$$(\mathcal{N}, e \circ \mu) \nvDash \psi,$$

so $(\mathcal{N}, e \circ v) \nvDash \varphi$. □

The theorem reveals that two isomorphic structures have identical truths; that is, they are homogeneous. Specifically, a formula satisfiable in one must be satisfiable in the other, and a formula valid in one must be valid in the other. Therefore, we cannot distinguish one structure from the other by merely looking at their respective truths in $\mathcal{L}$.

This consequence of isomorphism is significant. As we saw before, $\mathcal{L}$ cannot capture the mathematical meaning of its formulas, but is able to reveal their internal logic. Therefore, two isomorphic structures, although they might be two completely different mathematical contexts, since their truths are formally identical in $\mathcal{L}$, must have the same logical nature.

Meanwhile, when studying the truths of structures, it would be natural to have special interest in formulas that are always true within a structure (i.e., valid formulas). As a result, sentences, which, within a structure, can either be entirely true or entirely false, and which is satisfied either by all valuations or by none, deserve special attention. Thus, we arrive at the following definitions.

Definition 2.21. $\mathcal{L}_\mathcal{A}$-structures $\mathcal{M}$ and $\mathcal{N}$ are **elementary equivalent**, denoted as $\mathcal{M} \equiv \mathcal{N}$, iff for all $\mathcal{L}_\mathcal{A}$-sentences θ,

$$\mathcal{M} \vDash \theta \iff \mathcal{N} \vDash \theta.$$

Definition 2.22. Let $\mathcal{M}$ be an $\mathcal{L}$-structure. We define

$$Th(\mathcal{M}) = \{\theta | \theta \textit{ is a sentence, and } \mathcal{M} \vDash \theta\}.$$

We call $Th(\mathcal{M})$ the **theory** of $\mathcal{M}$.

Theorem 2.23. *Two isomorphic structures are elementary equivalent.*

Proof. By Theorem 2.20. □

Besides isomorphism, there are other structural relations. If we consider how we could attain one structure from another, these relations would naturally emerge.

Given a structure $\mathcal{M} = (M, I)$, we can expand M to M^*, so that we attain a new structure $\mathcal{M}^* = (M^*, I)$ with the same interpretation but a larger univserse, thereby attaining a new structural relation:

Definition 2.24. Let $\mathcal{M} = (M, I), \mathcal{N} = (N, J)$ be two $\mathcal{L}_\mathcal{A}$-structures. $\mathcal{M}$ is a **substructure** of $\mathcal{N}$, denoted as $\mathcal{M} \subseteq \mathcal{N}$, iff:

(1) $M \subseteq N$;

(2) if $c \in \mathcal{A}$ is a constant, then $I(c) = J(c)$;

(3) if $F \in \mathcal{A}$ is an n-variable function, then $I(F) = J(F) \restriction_{M^n}$;

(4) if $P \in \mathcal{A}$ is an n-variable relation, then $I(P) = J(P) \cap M^n$. We also call $\mathcal{M}$ a **reduction** of $\mathcal{N}$ and, equivalently, $\mathcal{N}$ an **expansion** of $\mathcal{M}$.

As we increase or decrease the universe of $\mathcal{M}$, we can intuitively say that $\mathcal{M}$ is increased or decreased, so that we attain a notion of the "size" of $\mathcal{M}$.

Definition 2.25. The size of a sturcture $\mathcal{M} = (M, I)$ is the cardinality of M, denoted as $||\mathcal{M}||$.

Combining the concept of homogeny and substructure, we may discover another structural relation:

Definition 2.26. Let $\mathcal{M} = (M, I), \mathcal{N} = (N, J)$ be two $\mathcal{L}_\mathcal{A}$-structures. $\mathcal{M}$ is an **elementary substructure** of $\mathcal{N}$, or, equivalently, $\mathcal{M}$ is an **elementary reduction** of $\mathcal{N}$ and $\mathcal{N}$ is an **elementary expansion** of $\mathcal{M}$, denoted as $\mathcal{M} \preceq \mathcal{N}$, iff $\mathcal{M}$ is a substructure of $\mathcal{N}$ and for all $\mathcal{L}_\mathcal{A}$-formulas φ and for all $\mathcal{M}$-valuations v,

$$(\mathcal{M}, v) \vDash \varphi \iff (\mathcal{N}, v) \vDash \varphi.$$

Theorem 2.27. *If $\mathcal{M}$ is an elementary substructure of $\mathcal{N}$, then $\mathcal{M} \equiv \mathcal{N}$.*

Proof. By Definition 2.26. □

Example 2.28. $(\mathbb{N}, <)$ is a substructure of $(\mathbb{Z}, <)$, and $(\mathbb{Z}, <)$ is a substructure of $(\mathbb{Q}, <)$, but both substructures are not elementary. However, $(\mathbb{Q}, <)$ is an elementary substructure of $(\mathbb{R}, <)$, that is, $(\mathbb{Q}, <) \preceq (\mathbb{R}, <)$.

Therefore, elementary substructures describe a relation where expanding or reducing the universe does not change a structure's logical nature. Then we may ask: by how much can we expand or reduce a structure without changing its logical nature? This question is answered by the following theorems.

Theorem 2.29. *(Downward Löwenheim-Skolem Theorem) Let $\mathcal{N} = (N, J)$ be an infinite $\mathcal{L}_\mathcal{A}$-structure, and $X \subseteq N$ be a countable set. Then $\mathcal{N}$ has a countable elementary reduction $\mathcal{M} = (M, I)$; that is, $X \subseteq M$, M is countable, and $\mathcal{M} \preceq \mathcal{N}$.*[5]

Theorem 2.30. *(Upward Löwenheim-Skolem Theorem) Let $\mathcal{L}_\mathcal{A}$ be countable,*[6] *$\mathcal{M} = (M, I)$ be an infinite $\mathcal{L}_\mathcal{A}$-structure, and κ be an infinite cardinal number. If there exists an injection from M to κ, then $\mathcal{M}$ must have a elementary expansion $\mathcal{N}$ where $||\mathcal{N}|| = \kappa$.*[7]

Theorem 2.29 shows that any infinite first order structure can be reduced to a countable structure while maintaining its logical nature, and Theorem 2.30 shows that any infinite first order structure can be arbitrarily expanded while maintaining its logical nature.

Moreover, given an $\mathcal{L}_\mathcal{A}$-structure $\mathcal{M} = (M, I)$, we can add new non-logical symbols to $\mathcal{A}$ to get a set of non-logical symbols $\mathcal{B}$ and expand I to I^* to include the interpretations of these new symbols. As a result, we get a $\mathcal{L}_\mathcal{B}$-structure $\mathcal{M}^* = (M, I^*)$, as well as a new structural relation, which is formally defined below:

Definition 2.31. Let $\mathcal{A}, \mathcal{B}$ be two sets of non-logical symbols, and $\mathcal{M} = (M, I), \mathcal{N} = (N, J)$ be $\mathcal{L}_\mathcal{A}, \mathcal{L}_\mathcal{B}$-structures respectively. $\mathcal{N}$ is an **increment** of $\mathcal{M}$, or, equivalently, $\mathcal{M}$ is a **shrinkage** of $\mathcal{N}$ iff

$$M = N, \mathcal{A} \subset \mathcal{B} \text{ and } I = J \restriction_\mathcal{A}.$$

[5] For proof, see page 117 of [2]. Also see Theorem Theorem 3.22 for another version of this theorem, whose proof is much simpler.

[6] See Definition 2.32.

[7] For proof, see page 183 of [2]. Also see Theorem Theorem 3.23 for another version of this theorem, which requires much simpler construction and proof.

Accordingly, $\mathcal{L}_\mathcal{A}$ is a **shrinkage** of $\mathcal{L}_\mathcal{B}$ and $\mathcal{L}_\mathcal{B}$ is an **increment** of $\mathcal{L}_\mathcal{A}$, denoted as $\mathcal{L}_\mathcal{A} \subset \mathcal{L}_\mathcal{B}$.

Note that the increment and shrinkage of a structure, like its expansion and reduction defined in Definition 2.24, are also about "increasing" and "decreasing" the structure. However, the former "increase" and "decrease" the structure by essentially increasing and decreasing the size of the *language* (that is, the number of non-logical symbols). Therefore, there was no ambiguity or arbitrariness in defining the "size" of a structure as the size of its universe in Definition 2.25. Nevertheless, to capture the idea of the size of a language, we give the following definition:

Definition 2.32. The size of a first order language $\mathcal{L}_\mathcal{A}$ is the cardinality of $\mathcal{A}$, denoted as $||\mathcal{L}_\mathcal{A}||$.

One typical way to attain a structure's increment is to add constants to the set of non-logical symbols.

Definition 2.33. Let $\mathcal{M} = (M, I)$ be an $\mathcal{L}_\mathcal{A}$-structure, $X \subseteq M$ be a nonempty set. We add a new constant c_a to $\mathcal{A}$ for every $a \in X$; that is, we let

$$\mathcal{A}_X = \mathcal{A} \cup \{c_a | a \in X\},$$

so that we get $\mathcal{L}_{\mathcal{A}_X}$, an increment of $\mathcal{L}_\mathcal{A}$, and $\mathcal{M}_X = (M, I_X)$, an increment of $\mathcal{M}$ where $I_X(c_a) = a$ for all $a \in M$. We call $\mathcal{L}_{\mathcal{A}_X}$ the **X-constant increment** of $\mathcal{L}_\mathcal{A}$, and $\mathcal{M}_X$ the **X-constant increment** of $\mathcal{M}$.

Constant increments allow a language to refer to more elements in the universe. As a result, the set of truths of a structure, after constant increment, will be expanded to include true statements about these elements, so that their properties can be directly delineated by the language. As a result, the structure becomes more distinct from other structures in its truths. This intuitive idea can be illustrated by the following theorem:

Theorem 2.34. *Let $\mathcal{M} = (M, I), \mathcal{N} = (N, J)$ be two $\mathcal{L}_\mathcal{A}$-structures, where $\mathcal{M} \subseteq \mathcal{N}$. For every $a \in M$ we introduce a new constant c_a to $\mathcal{L}_\mathcal{A}$. We thus form the language $\mathcal{L}_{\mathcal{A}_M}$, where*

$$\mathcal{A}_M = \mathcal{A} \cup \{c_a | a \in M\}.$$

Therefore, we get constant increments of $\mathcal{M}$ and $\mathcal{N}$: $\mathcal{M}^* = (M, I^*), \mathcal{N}^* = (N, J^*)$, where

$$I^*(c_a) = J^*(c_a) = a, a \in M.$$

Then

$$\mathcal{M} \preceq \mathcal{N} \iff \mathcal{M}^* \equiv \mathcal{N}^*.$$ [8]

As we have seen, elementary expansion and reduction imply greater logical similarity between two structures than elementary equivalence does; therefore, constant increment, changing the structual relation from the former to the latter, makes two structures more differentiated from each other.

However, we should note that a change in relation between two structures after constant increment does not imply a change in the *mathematical contexts* they represent. In Example 2.28 we saw that $(\mathbb{Q}, <) \preceq (\mathbb{R}, <)$. If we apply the constant increment in Theorem 2.34 to $(\mathbb{Q}, <)$ and $(\mathbb{R}, <)$, their structural relation changes as the language's power to differentiate them increases, but it would be absurd to

[8] See page 113 of [2].

say that the logical relation between ordering on $\mathbb{Q}$ and ordering on $\mathbb{R}$ also changes. Therefore, an $\mathcal{L}$-structure, representing a mathematical context, might not be able to fully capture its logical nature. This shortcoming, by the above theorem, is partly due to the language's inability to refer to every element in the universe.

3. First Order System

We can use $\mathcal{L}$-formulas to represent mathematical statements. However, we still do not have the means to derive one $\mathcal{L}$-formula from another, while in mathematics, this is exactly the function of logic. Therefore, to proceed further in our inquiry, we need to build a deduction system for $\mathcal{L}$. Such system should consists of 1) a set of axioms, which provides a starting point for any deduction, and 2) rules of deduction, which allow us to deduce one formula from another. Because of the conrrespondence between mathematical statements and $\mathcal{L}$-formulas, by studying this system we can understand mathematical deductions more deeply.

However, it is not the case that any system with a set of $\mathcal{L}$-formulas as axioms and some rules of deduction is desirable. In mathematics we want and only want to deduce true statements. Therefore, we expect the system to be powerful enough to derive as many true statements as possible (i.e., it should be **complete**); meanwhile, it should not be too powerful, as from true statements it should be able to derive true statements only (i.e., it should be **sound**), and it should not be able to derive two paradoxical statements (i.e., it should be **consistent**).

We now define a first order deduction system with seven axioms and one rule of deduction. We will later show that this system is indeed our desirable system.

Definition 3.1. (Logical Axioms) The set of first order logical axioms, $\mathbb{L}$, is the minimal set of $\mathcal{L}$-formulas satisfying the following:

(1) Let $\varphi_1, \varphi_2, \varphi_3$ be $\mathcal{L}$-formulas. Then the following formulas are in $\mathbb{L}$:
  - (a) $((\varphi_1 \to (\varphi_2 \to \varphi_3)) \to ((\varphi_1 \to \varphi_2) \to (\varphi_1 \to \varphi_3)))$;
  - (b) $(\varphi_1 \to \varphi_1)$;
  - (c) $(\varphi_1 \to (\varphi_2 \to \varphi_1))$;
  - (d) $(\varphi_1 \to ((\neg\varphi_1) \to \varphi_2))$;
  - (e) $(((\neg\varphi_1) \to \varphi_1) \to \varphi_1)$;
  - (f) $((\neg\varphi_1) \to (\varphi_1 \to \varphi_2))$;
  - (g) $(\varphi_1 \to ((\neg\varphi_2) \to (\neg(\varphi_1 \to \varphi_2))))$.

(2) Let φ be a formula, τ be a term. If τ can substitute variable x in φ, then
$$((\forall x\varphi) \to \varphi(x;\tau)) \in \mathbb{L}.$$

(3) Let φ_1, φ_2 be $\mathcal{L}$-formulas. Then
$$((\forall x(\varphi_1 \to \varphi_2)) \to ((\forall x\varphi_1) \to (\forall x\varphi_2))) \in \mathbb{L}.$$

(4) Let φ be a formula. If x is not a free variable in φ, then
$$(\varphi \to (\forall x\varphi)) \in \mathbb{L}.$$

(5) If $\varphi \in \mathbb{L}$, then $(\forall x\varphi) \in \mathbb{L}$.

(6) For every variable x, $(x \equiv x) \in \mathbb{L}$.

(7) Let φ_1 and φ_2 be two formulas, and x_j can substitute x_i in φ_1 and φ_2. If the two formulas formed by substituting x_j for every occurrence of x_i in φ_1 and φ_2 respectively are the same, then
$$((x_j \equiv x_i) \to (\varphi_1 \to \varphi_2)) \in \mathbb{L}.$$

Definition 3.2. (Rule of Deduction) We can deduce formula ψ from formulas φ and $(\varphi \to \psi)$.

Note that the system established above uses purely formal transformation to represent the logical deduction from one statement to another. This reveals an important idea we saw in the previous section: logic is represented formally in $\mathcal{L}$.

Meanwhile, as we are now able to derive formulas from formulas, the following definitions become natural:

Definition 3.3. Let Γ be a set of $\mathcal{L}$-formulas. A finite sequence of $\mathcal{L}$-formulas $\langle \varphi_1, \ldots \varphi_n \rangle$, where $n \geq 1$, is a **proof** from Γ to ψ iff such sequence satisfies the following:
(1) $\varphi_1 \in \Gamma \cup \mathbb{L}$;
(2) $\varphi_n = \psi$;
(3) for every i, where $1 \leq i \leq n$, either
  (a) $\varphi_i \in \Gamma \cup \mathbb{L}$, or
  (b) there exist j, k, where $1 \leq j, k < i$, such that φ_j is $\varphi_k \to \varphi_i$.

Definition 3.4. Let Γ be a set of $\mathcal{L}$-formulas, φ be an $\mathcal{L}$-formula. φ is **provable** from Γ, or, equivalently, φ is a **theorem** of Γ, denoted as

$$\Gamma \vdash \varphi,$$

iff there exists a proof from Γ to φ. It follows that $\varphi \in \mathbb{L}$ iff $\vdash \varphi$.

Example 3.5. We show $\vdash ((\neg(\neg\varphi)) \to \varphi)$ as an example of logical proof.

Proof. (1) $\vdash ((\neg(\neg\varphi)) \to ((\neg\varphi) \to \varphi))$, by Definition 3.1(1(f));
(2) $\vdash (((\neg\varphi) \to \varphi) \to \varphi)$, by Definition 3.1(1(e));
(3) $\vdash ((((\neg\varphi) \to \varphi) \to \varphi) \to ((\neg(\neg\varphi)) \to (((\neg\varphi) \to \varphi) \to \varphi)))$, by Definition 3.1(1(c));
(4) $\vdash ((\neg(\neg\varphi)) \to (((\neg\varphi) \to \varphi) \to \varphi))$, by applying the Rule of Deduction to (2) and (3);
(5) $\vdash (((\neg(\neg\varphi)) \to (((\neg\varphi) \to \varphi) \to \varphi)) \to ((\neg(\neg\varphi)) \to ((\neg\varphi) \to \varphi) \to ((\neg(\neg\varphi)) \to \varphi)))$, by Definition 3.1(1(a));
(6) $\vdash ((\neg(\neg\varphi)) \to ((\neg\varphi) \to \varphi) \to ((\neg(\neg\varphi)) \to \varphi))$, by applying the Rule of Deduction to (4) and (5);
(7) $\vdash ((\neg(\neg\varphi)) \to \varphi)$, by applying the Rule of Deduction to (1) and (6). $\square$

The following theorem provides a shortcut for many logical proofs.

Theorem 3.6. *(Deduction Theorem) Let Γ be a set of $\mathcal{L}$-formulas, and φ_1, φ_2 be two formulas. Then*

$$\Gamma \cup \varphi_1 \vdash \varphi_2 \iff \Gamma \vdash (\varphi_1 \to \varphi_2).$$ [9]

With the definitions and notations just introduced, we may now formally define **consistency**.

Definition 3.7. Let Γ be a set of $\mathcal{L}$-formulas. Γ is **consistent** iff for any $\mathcal{L}$-formula φ, if $\Gamma \vdash \varphi$, then $\Gamma \nvdash (\neg\varphi)$.

From this definition we may see a consequence of our system that connects it to the normal mathematical reasoning:

[9] For proof, see page 130 of [2].

Theorem 3.8. *Let Γ be a set of $\mathcal{L}$-formulas, φ be an $\mathcal{L}$-formula, then*
(1) $\Gamma \cup \{(\neg\varphi)\}$ is not consistent iff $\Gamma \vdash \varphi$.
(2) $\Gamma \cup \{\varphi\}$ is not consistent iff $\Gamma \vdash (\neg\varphi)$.[10]

Therefore, given a set of hypotheses Γ, if adding a statement to Γ makes it inconsistent, we are allowed to deduce the negation of the statement. This reveals that proof by contradiction holds in our system.

However, that our system is able to deduce proof by contradiction does not mean that proof by contradiction is a logical truth. In fact, since all theorems of the system are merely results of formal transformation, there is no reason, at this point, to claim the truth of any of them, and, moreover, if we cannot ensure their truth, the system would be unworthy for any further study. Therefore, a proof of its **soundness** becomes urgent at this point. We begin the proof with a lemma:

Lemma 3.9. *Every axiom in $\mathbb{L}$ is universally valid.*

Proof. Suppose $\varphi \in \mathbb{L}$. Let $\mathcal{M}$ be an $\mathcal{L}$-structure and v be a valuation. We want to show

$$(\mathcal{M}, v) \vDash \varphi.$$

By Definition 3.1 there are 7 possible scenarios.

(1) φ is a tautology in propositional logic. Then φ is also a tautology in $\mathcal{L}$.

(2) φ is $((\forall x\psi) \to \psi(x;\tau))$, where τ can substitute x_i in ψ.

If $(\mathcal{M}, v) \nvDash (\forall x\psi)$, we have $(\mathcal{M}, v) \vDash \varphi$ by definition.

If $(\mathcal{M}, v) \vDash (\forall x\psi)$, then for any $\mathcal{M}$-valuation μ, if $\mu \equiv_{(\forall x\psi)} v$, then $(\mathcal{M}, \mu) \vDash \psi$. Consider the valuation μ^*, where $\mu^*(x) = \bar{v}(\tau), \mu^* \equiv_{(\forall x\psi)} v$. We have

$$(\mathcal{M}, \mu^*) \vDash \psi.$$

By the Replacement Theorem, $(\mathcal{M}, v) \vDash \psi(x;\tau)$. Therefore, $(\mathcal{M}, v) \vDash \varphi$.

(3) φ is $((\forall x(\psi_1 \to \psi_2)) \to ((\forall x\psi_1) \to (\forall x\psi_2)))$.

Suppose $(\mathcal{M}, v) \vDash (\forall x(\psi_1 \to \psi_2)), (\mathcal{M}, v) \vDash (\forall x\psi_1)$. Then for any valuation μ, if

$$\mu = v \bmod ((\forall x(\psi_1 \to \psi_2))),$$

we have $(\mathcal{M}, \mu) \vDash \psi_2$.

Now let μ be an $\mathcal{M}$-valuation, and $\mu = v \bmod ((\forall x\psi_2))$. We want to prove $(\mathcal{M}, \mu) \vDash \psi_2$. We define a valuation μ^*, where $\mu^*(x_i) = \mu(x_i)$ for any free variable x_i in ψ_2 and $\mu^*(x_j) = v(x_j)$ for any free variable x_j in $(\forall x(\psi_1 \to \psi_2))$ that is not free in ψ_2. Then $\mu^* = v \bmod (\forall x(\psi_1 \to \psi_2))$, so $(\mathcal{M}, \mu^*) \vDash \psi_2$ and therefore $(\mathcal{M}, \mu) \vDash \psi_2$.

(4) φ is $(\psi \to (\forall x\psi))$, where x is not a free variable in ψ. Suppose $(\mathcal{M}, v) \vDash \psi$. Then for all valuations μ, if $\mu = v \bmod (\psi)$, $(\mathcal{M}, \mu) \vDash \psi$. Since x is not a free variable in ψ, ψ and $(\forall x\psi)$ have the same free variables. Therefore, for any valuation μ, if $\mu = v \bmod ((\forall x\psi))$, $(\mathcal{M}, \mu) \vDash \psi$, so that $(\mathcal{M}, v) \vDash \psi$.

(5) φ is $(\forall x\psi)$, where $\psi \in \mathbb{L}$. By induction hypothesis, for any $\mathcal{M}$-valuation μ, $(\mathcal{M}, \mu) \vDash \psi$, so $(\mathcal{M}, v) \vDash (\forall x\psi)$.

(6) φ is $(x \equiv x)$. Then φ is valid in $\mathcal{M}$ by definition.

(7) φ is $((x_i \equiv x_j) \to (\psi_1 \to \psi_2))$, where x_j can replace x_i in ψ_1 and ψ_2, and $\psi_1(x_i; x_j)$ and $\psi_2(x_i; x_j)$ are identical.

[10] For proof, see page 131 of [2].

Suppose $(\mathcal{M}, v) \vDash (x_i \equiv x_j)$, and $(\mathcal{M}, v) \vDash \psi_1$. Then $v(x_j) = v(x_i)$. By the Replacement Theorem, $(\mathcal{M}, v) \vDash \psi_1(x_i; x_j)$. Therefore, $(\mathcal{M}, v) \vDash \psi_2(x_i; x_j)$. We apply the Replacement Theorem once again and get $(\mathcal{M}, v) \vDash \psi_2$. □

Theorem 3.10. *(Soundness) Let Γ be a set of $\mathcal{L}$-formulas, and φ be an $\mathcal{L}$-formula. If $\Gamma \vdash \varphi$, then $\Gamma \vDash \varphi$.*

Proof. Let $(\mathcal{M}, v) \vDash \Gamma$, and $\langle \varphi_1, \dots, \varphi_n \rangle$ be a proof from Γ to φ. We prove that for any $1 \le i \le n$, $(\mathcal{M}, v) \vDash \varphi_i$ by induction.

Suppose $\varphi_i \in \mathbb{L} \cup \Gamma$. If $\varphi_i \in \mathbb{L}$, by Lemma 3.9 $(\mathcal{M}, v) \vDash \varphi_i$. If $\varphi_i \in \Gamma$, then $(\mathcal{M}, v) \vDash \varphi_i$ since $(\mathcal{M}, v) \vDash \Gamma$. Note that if $i = 1$ we must have $\varphi_i \in \mathbb{L} \cup \Gamma$.

Suppose there exist $j, k < i, j \ne k$ such that φ_k is $(\varphi_j \to \varphi_i)$. By induction hypothesis, $(\mathcal{M}, v) \vDash \varphi_j, (\mathcal{M}, v) \vDash \varphi_k$. Then $(\mathcal{M}, v) \vDash \varphi_i$ by definition. □

Therefore, if Γ is true and we are able to deduce φ from Γ, φ is also true; that is, our system from true statements can deduce true statements only. If we let $\Gamma = \emptyset$, then by the theorem $\vdash \varphi$ implies $\vDash \varphi$. Therefore, all axioms and theorems of the system defined in Definition 3.1 are universally valid.

It might be surprising at first that theorems, results of *formal* transformation, can be equated with truth, which is a *semantical* concept. However, we shall recall that $\mathcal{L}$ itself bears logical meanings. In fact, the connection between truths and theorems is due to 1) the logical meanings of $\mathcal{L}$, and 2) the axioms and rules of deduction prescribed by the system. Specifically, each transformation of $\mathcal{L}$-formulas (or, equivalently, each axiom and rule of deduction), by the logical meanings born by their symbols, correspond to certain normal logical reasoning, and, therefore, Theorem 3.10 essentially states that such reasoning logically sound — that is, it infers truths from truths, thereby valid formulas from valid formulas. Indeed, for instance, the Rule of Deduction is essentially a formalization of syllogism, the well-known classical reasoning with indisputable validity.

More insights can be attained if we consider the following consequence of Theorem 3.10:

Theorem 3.11. *Let Γ be a set of $\mathcal{L}$-formulas. If Γ is satisfiable, then Γ is consistent.*

Proof. We prove by contradiction. Suppose there exists a set Γ of formulas that is satisfiable but not consistent. Then there exists a formula φ such that $\Gamma \vdash \varphi, \Gamma \vdash (\neg\varphi)$. By Theorem 3.10, $\Gamma \vDash \varphi, \Gamma \vDash (\neg\varphi)$. Since Γ is satisfiable, let it be satisfied in structure $\mathcal{M}$ by valuation v, then $(\mathcal{M}, v) \vDash \varphi, (\mathcal{M}, v) \vDash (\neg\varphi)$, which is a contradiction. □

Therefore, the truths within any mathematical context *that can be represented by* $\mathcal{L}$,[11] since they are satisfiable, must be consistent. The consistency of mathematics has been partially proven.

In fact, what characterize an ideal deduction system should be more than soundness. Consider a somewhat extreme case. If all theorems of a deduction system are true, while the system is only able to deduce one theorem, it would also be pointless to study the system. Therefore, ideally, a deduction system should be powerful enough to deduce all true statements. This property is called **completeness**. We will now show that our deduction system is indeed complete:

[11] This condition is crucial, as not all mathematical statements can be represented by $\mathcal{L}$. See Theorem 3.21.

Theorem 3.12. *(Completeness) Let Γ be a set of $\mathcal{L}$-sentences and φ be an $\mathcal{L}$-sentence. If $\Gamma \vdash \varphi$, then $\Gamma \vDash \varphi$.*

We begin the proof of this theorem by introducing the following definitions.

Definition 3.13. Let Γ be a set of $\mathcal{L}$-sentences, C be a set of $\mathcal{L}$-constants. We say that C is a **set of witnesses** for Γ in $\mathcal{L}$ iff for every $\mathcal{L}$-formula φ with at most one free variable, say x, there is $c \in C$ such that $\Gamma \vdash (\exists x \varphi) \to \varphi(x; c)$. We say that Γ has **witnesses** in $\mathcal{L}$ iff Γ has some set C of witnesses in $\mathcal{L}$.

Definition 3.14. Let Γ be a consistent set of formulas. Γ is **maximal consistent** iff for any $\mathcal{L}$-formula φ, either $\varphi \in \Gamma$, or $\Gamma \cup \{\varphi\}$ is not consistent.

Next, we prove the following theorem, from which we will finally deduce Theorem 3.12.

Theorem 3.15. *Let Γ be a set of $\mathcal{L}$-sentences. If Γ is consistent, then Γ is satisfiable.*

We divide the proof of Theorem 3.15 into two parts. In the first part, we show that every consistent set of sentences can be extended to a consistent set that has witnesses. Then we show that any consistent set of sentences that has a set of witnesses has a model, thereby being satisfiable.

Lemma 3.16. *Let Γ be a consistent set of $\mathcal{L}_{\mathcal{A}}$-sentences. Let C be a set of new constant symbols where $|C| = ||\mathcal{L}_{\mathcal{A}}||$, and let $\mathcal{L}_{\mathcal{B}}$ be the constant increment of $\mathcal{L}_{\mathcal{A}}$ such that $\mathcal{B} = \mathcal{A} \cup C$. Then Γ can be extended to a consistent set Γ^* of $\mathcal{L}_{\mathcal{B}}$-sentences that has C as a set of witnesses in $\mathcal{L}_{\mathcal{B}}$.*

Proof. Let $\alpha = ||\mathcal{L}_{\mathcal{A}}||$. For every $\beta < \alpha$, let c_β be a constant which does not occur in $\mathcal{L}_{\mathcal{A}}$, such that $c_\beta \neq c_\gamma$ if $\beta < \gamma\alpha$. Let

$$C = \{c_\beta : \beta < \alpha\},$$

and

$$\mathcal{B} = \mathcal{A} \cup C.$$

Clearly, $||\mathcal{L}_{\mathcal{B}}|| = \alpha$, so we may enumerate all $\mathcal{L}_{\mathcal{B}}$-formulas with at most one free variable in a sequence $\varphi_n, n < \alpha$. We now define an increasing sequence of sets of $\mathcal{L}_{\mathcal{B}}$-sentences:

$$\Gamma = \Gamma_0 \subset \Gamma_1 \subset \cdots \subset \Gamma_n \subset \dots$$

where $n < \alpha$, and a sequence of constants $d_n, n < a$ from C such that:

(1) each Γ_i is consistent in $\mathcal{L}_{\mathcal{B}}$;

(2) if $i = j+1$, then $\Gamma_i = \Gamma_j \cup \{(\exists x_j)\varphi_j \to \varphi_j(d_j)\}$, where x_j is the free variable in φ_m if it has one, otherwise $x_j = v_0$;

(3) if n is a limit ordinal different from 0, then $\Gamma_n = \bigcup_{m<n} \Gamma_m$.

Suppose Γ_n has been defined. Note that the number of sentences in Γ_n that are not $\mathcal{L}_{\mathcal{A}}$-sentences is smaller than α. Furthermore, each such sentence contains a finite number of constants from C. Therefore, let d_n be the first element of C which has not yet occurred in Γ_n. For instance, $d_0 = c_0$. We show that

$$\Gamma_{n+1} = \Gamma_n \cup \{(\exists x_n \varphi_n) \to \varphi_n(d_n)\}$$

is consistent. If this were not the case, then

$$\Gamma_n \vdash \neg((\exists x_n \varphi_n) \to \varphi_n(d_n)).$$

By propositional logic,

$$\Gamma_n \vdash ((\exists x_n \varphi_n) \wedge (\neg\varphi(d_n))).$$

As d_n does not occur in T_n, we have

$$\Gamma_n \vdash (\forall x_n((\exists x_n \varphi_n) \wedge (\neg\varphi_n(x_n)))),$$

$$\Gamma_n \vdash ((\exists x_n \varphi_n) \wedge (\neg(\exists x_n \varphi_n)),$$

which contradicts the consistency of Γ_n. If m is a nonzero limit ordinal, and each member of the increasing chain $\Gamma_n, n < m$ is consistent, then $\Gamma_m = \bigcup_{n<m} \Gamma_n$ is consistent. This completes the induction.

Now let $\Gamma^* = \bigcup_{m<\alpha} \Gamma_m$. Then Γ^* is consistent and $\Gamma^* \supset \Gamma$. Suppose φ is an $\mathcal{L}_\mathcal{B}$-formula with at most one free variable x. Then we may suppose $\varphi = \varphi_m$ and $x = x_m$ for some $m < \alpha$. Whence the sentence

$$((\exists x_m \varphi_m) \to \varphi_m(d_m))$$

belongs to Γ_{m+1} and so to Γ^*. □

Lemma 3.17. *Let Γ be a consistent set of $\mathcal{L}$-sentences and C be a set of witnesses for Γ in $\mathcal{L}$. Then Γ has a model $\mathcal{M} = (M, I)$ such that every element of M is an interpretation of a constant $c \in C$.*

Proof. Note that if a set of sentences Γ has witnesses of C in $\mathcal{L}$, for every set of sentences $\Sigma \supset \Gamma$: (1) Σ also has C as witnesses; (2) if Σ has a model $\mathcal{M}$, then $\mathcal{M}$ is also a model of Γ. So we may as well assume that Γ is maximal consistent in $\mathcal{L}$.

We let C be a set of witnesses for Γ. To show that Γ has a model, we will first construct a structure $\mathcal{M}$, and then show that it is a model of Γ.

Part 1.

To define a structure, we need to first define its universe.

For two constants $c, d \in C$, define

$$c \sim d \text{ iff } (c \equiv d) \in \Gamma.$$

Because Γ is maximal consistent, we see that for $c, d, e \in C$,

$$c \sim c;$$
$$\text{if } c \sim d \text{ and } d \sim e, \text{ then } c \sim e;$$
$$\text{if } c \sim d \text{ then } d \sim c.$$

So $\sim$ is an equivalence relation on C. For each $c \in C$, let

$$\tilde{c} = \{d \in C : d \sim c\}$$

be the equivalence class of c. We let $M = \{\tilde{c}{:}c \in C\}$ be the set of all equivalence classes in C. We let M be the universe of our structure.

Then we define the interpretation I of $\mathcal{M}$.

(Constant) Let d be an $\mathcal{L}$-constant. Since

$$\vdash (\exists x_0(d \equiv x_0)),$$

$(\exists x_0(d \equiv x_0)) \in \Gamma$. Because Γ has witnesses of C, there is a constant $c \in C$ such that

$$(d \equiv c) \in \Gamma.$$

By Definition 3.1(7),

$$\vdash (d \equiv c \wedge d \equiv c' \to c \equiv c').$$

Thus, for every constant c, its equivalence class $\tilde{c}$ is well-defined, so we can comfortably let c be interpreted in $\mathcal{M}$ as $\tilde{c} \in M$.

Note that every element in M is the equivalence class of some $c \in C$; that is, every element in M is the interpretation of some constant $c \in C$.

(Relation) For each n-variable $\mathcal{L}$-relation P, we define an n-variable relation P' on the set C such that for all $c_1, \dots, c_n \in C$,

$$P'(c_1, \dots, c_n) \text{ iff } P(c_1, \dots, c_n) \in \Gamma.$$

By Definition 3.1(7),

$$\vdash P'(c_1, \dots, c_n) \wedge c_1 \equiv d_1 \wedge \cdots \wedge c_n \equiv d_n \to P'(d_1, \dots, d_n).$$

Therefore, P' only depends on the interpretation of its variables, so it is well-defined.

We thus for every $\mathcal{L}$-relation P define a relation $\tilde{P}$ on M^n such that

$$\tilde{P}(\tilde{c_1}, \dots, \tilde{c_n}) \text{ iff } P(c_1, \dots, c_n) \in \Gamma.$$

We let $\tilde{P}$ be the interpretation of P in $\mathcal{M}$.

(Function) Let F be any n-variable $\mathcal{L}$-function, and let $c_1 \dots, c_n \in C$. Then

$$(\exists x_0(F(c_1, \dots, c_n) \equiv x_0)) \in \Gamma.$$

Because Γ has witnesses, there is a constant $c \in C$ such that

$$(F(c_1, \dots, c_n) \equiv c) \in \Gamma.$$

By Definition 3.1(7),

$$\vdash (F(c_1, \dots c_n) \equiv c \wedge c_1 \equiv d_1 \wedge \cdots \wedge c_n \equiv d_n \wedge c \equiv d) \to F(d_1, \dots, d_n) \equiv d.$$

Therefore, the value of F only depends on the interpretation of its variables, so it is well-defined.

We thus for every $\mathcal{L}$-function F define a function $\tilde{F}$ on M such that $\tilde{F}(\tilde{c_1}, \dots, \tilde{c_n}) = \tilde{c}$ iff $(F(c_1, \dots, c_n) \equiv c) \in \Gamma$. We let $\tilde{F}$ be the interpretation of F in $\mathcal{M}$.

We have now completed the definition of $\mathcal{M}$.

Part 2.

We now prove that for any sentence θ,

$$\mathcal{M} \vDash \theta \iff \theta \in \Gamma.$$

We prove by induction on the length of θ:

By the definitions in part 1, for every $\mathcal{L}$-term τ without free variable and for every constant $c \in C$,

$$\mathcal{M} \vDash (\tau \equiv c) \iff (\tau \equiv c) \in \Gamma.$$

Since C is a set of witnesses for Γ:

(1) for any two terms τ_1, τ_2 without free variable,

$$\mathcal{M} \vDash (\tau_1 \equiv \tau_2) \iff (\tau_1 \equiv \tau_2) \in \Gamma;$$

(2) for any atomic formula $P(\tau_1, \dots, \tau_n)$ containing no free variables,

$$\mathcal{M} \vDash P(\tau_1, \dots, \tau_n) \iff P(\tau_1, \dots, \tau_n) \in \Gamma.$$

From the scenarios above it is not difficult to see that for any sentences φ, ψ,

$$\mathcal{M} \vDash (\neg\varphi) \iff (\neg\varphi) \in \Gamma,$$

and

$$\mathcal{M} \vDash (\psi \wedge \varphi) \iff (\psi \wedge \varphi) \in \Gamma.$$

Suppose θ is $(\exists x\varphi)$. If $\mathcal{M} \vDash \theta$, then for some $\tilde{c} \in M$, $\mathcal{M} \vDash \varphi[\tilde{c}]$. This means that $\mathcal{M} \vDash \varphi(x;c)$, where $\varphi(x;c)$. So $\varphi(x;c) \in \Gamma$. Because $\vdash \varphi(x;c) \to (\exists x\varphi)$, $\theta \in \Gamma$.

On the other hand, if $\theta \in \Gamma$, then because Γ has witnesses, there exists a constant $c \in C$ such that

$$\Gamma \vdash (\exists x\varphi) \to \varphi(c).$$

Since Γ is maximal consistent, $\varphi(x;c) \in \Gamma$, so $\mathcal{M} \vDash \varphi(x;c)$. This gives $\mathcal{M} \vDash \varphi[\tilde{c}]$ and $\mathcal{M} \vDash \theta$. This shows that $\mathcal{M}$ is a model of Γ. □

Now the proof of Theorem 3.15 becomes straightforward:

Proof. Suppose Γ is consistent. By Lemma 3.16, we consider the constant increments Γ^* of Γ and $\mathcal{L}^*$ of $\mathcal{L}$, so that Γ^* has witnesses in $\mathcal{L}^*$. By Lemma 3.17, let $\mathcal{M}^* = (M^*, I^*)$ be a model of Γ^*. $\mathcal{M}^*$ is a model for the incremented language $\mathcal{L}^*$, so let $\mathcal{M} = (M, I)$ be the $\mathcal{L}$-structure that is the shrinkage of $\mathcal{M}^*$ to $\mathcal{L}$. Because sentences in Γ do not involve $\mathcal{L}^*$-constants that are not in $\mathcal{L}$, $\mathcal{M}$ is a model of Γ. Therefore, Γ is satisfiable. □

Combining Theorem 3.11 and Theorem 3.15, we attain the following:

Theorem 3.18. *Let Γ be a set of $\mathcal{L}$-sentences. Γ is consistent iff Γ is satisfiable.*

Consistency and satisfiability are equivalent in $\mathcal{L}$. As a result, the necessary and sufficient condition for a set of $\mathcal{L}$-sentences to be true within some mathematical context is that it has no logical contradiction. That is, every set of consistent statements that can be represented by $\mathcal{L}$ can find a place among the truths of mathematics. In this sense, logical consistency is the sole requirement for any set of statements to be accepted by mathematics. This partially implies that logic is the foundation of mathematics.

Finally, we are able to prove Theorem 3.12:

Proof. We prove by contradiction. Suppose $\Gamma \vDash \varphi$. If $\Gamma \nvdash \varphi$, by Theorem 3.8(1), $\Gamma \cup \{(\neg\varphi)\}$ is consistent, and by Theorem 3.18 it is satisfiable.

Let $\mathcal{M}$ be a structure that satisfies $\Gamma \cup \{(\neg\varphi)\}$. Then

$$\mathcal{M} \vDash \Gamma \cup \{(\neg\varphi)\}.$$

Then $\mathcal{M} \vDash \Gamma$, so $\mathcal{M} \vDash \varphi$. There is a contradiction. □

Therefore, if from Γ we are able to deduce φ, then if Γ is true, φ must also be true. If we let $\Gamma = \emptyset$, then $\vDash \varphi$ implies $\vdash \varphi$. Therefore, every true $\mathcal{L}$-sentence can be deduced by our system.

Moreover, Theorem 3.12, like Theorem 3.10, also implies a connection between theorems and truths, and such connection can also be explained by: 1) the logical meanings contained in $\mathcal{L}$, and 2) the axioms and rules of deduction prescribed by the system. Specifically, given a set of sentences, the axioms and rules of deduction prescribe the possible formal transformations from the set, each of which correspond to a logical reasoning. And Theorem 3.12 essentially states that every correct logical reasoning — which infers truths from truths — applicable to a set of statements corresponds to some transformation, or, equivalently, some axiom or rule of deduction, of our system. In other words, each normal, logical deduction in mathematics is represented some formal deduction. Therefore, our deduction system is powerful enough.

Combining Theorem 3.10 and Theorem 3.12, we have:

Theorem 3.19. *(Completeness Theorem) Let Γ be a set of $\mathcal{L}$-sentences, φ be an $\mathcal{L}$-sentence. Then*

$$\Gamma \vdash \varphi \iff \Gamma \vDash \varphi.$$

Therefore, completeness and soundness are equivalent for $\mathcal{L}$, and for simplicity, I call Theorem 3.19 *the* Completeness Theorem.

The Completeness Theorem, by equating theorems and truths, seems to imply a unification of form and intension in $\mathcal{L}$'s representation of mathematical statements. However, we shall not forget the distinction and gap between them that has been constantly emphasized. Indeed, form and intension, by definition, cannot be one entity, and there shall be no necessary relation between them. As we saw earlier, the connection between theorems and truths (and, thus, of form and intension) is merely the result of 1) $\mathcal{L}$'s logical meanings and 2) the axioms and rule of deduction prescribed. Therefore, to be precise, what Theorem 3.18 and the Completeness Theorem uncover is not a unification, but a *parallelism*, between theorems and truths, between form and intension, in $\mathcal{L}$.

The completeness of our deduction system has many deep and significant implications, and one of them is the Compactness Theorem.

Theorem 3.20. *(Compactness Theorem) A set of $\mathcal{L}$-sentences Γ is consistent iff every finite subset of Γ is consistent.*

Proof. ($\Rightarrow$) Supppose Γ is consistent. Then it is satisfiable by Theorem 3.18. Let $(\mathcal{M}, v) \vDash \Gamma$.

Let Γ_0 be a finite subset of Γ, then $(\mathcal{M}, v) \vDash \Gamma_0$, so that Γ_0 is satisfiable. By Theorem 3.18, Γ_0 is consistent.

($\Leftarrow$) We prove by contradiction. Suppose every finite subset of Γ is consistent. If Γ is not consistent, let θ be a sentence such that $\Gamma \vdash \theta$ and $\Gamma \vdash (\neg\theta)$. We let $\langle \alpha_1, \ldots, \alpha_n \rangle, \langle \beta_1, \ldots, \beta_m \rangle$ be the proofs from Γ to θ and to $(\neg\theta)$ respectively.

We let $\Gamma_0 = \{\alpha_i, \beta_j | \alpha_i \in \Gamma, \beta_j \in \Gamma\}$. Since Γ_0 is a finite subset of Γ, it is consistent. There is a contradiction. $\square$

The Compactness Theorem is significant as it has both wide and deep applications that allow us to attain great insight into both $\mathcal{L}$ and mathematics. The following theorem is an example.

Theorem 3.21. *If a set of sentences Γ has arbitrarily large finite models, then it has an infinite model.*

Proof. Let Γ be a set of $\mathcal{L}_{\mathcal{A}}$-sentences with arbitrarily large finite models. Consider the increment $\mathcal{B} = \mathcal{A} \cup \{c_n : n \in \omega\}$, where c_n is a list of distinct constant symbols not in $\mathcal{A}$. Consider the set

$$\Sigma = \Gamma \cup \{(\neg(c_n \equiv c_m)) : n < m < \omega\}$$

of $\mathcal{L}_{\mathcal{B}}$-sentences. Any finite subset Σ' of Σ involves at most the constants $c_0, \ldots, c_m$, where $m < \omega$. Let $\mathcal{M} = (M, I)$ be a model of Γ with at least $m+1$ elements, and let $X = \{a_0, \ldots, a_m\}$ be a set of $m+1$ distinct elements in M. Then the X-constant increment of $\mathcal{M}$ is a model of Σ'. Therefore, by the Compactness Theorem, Σ has a model. The shrinkage of this model to $\mathcal{L}_{\mathcal{A}}$ is an infinite model of Γ. $\square$

Therefore, any set of $\mathcal{L}$-sentences is unable to distinguish between infinity and finity. That is, infinity or finity cannot be defined or delineated by $\mathcal{L}$-formulas.

Therefore, $\mathcal{L}$ is unable to represent all mathematical statements, and through $\mathcal{L}$ we can at best attain truths about a limited part of logic and mathematics. Indeed, $\mathcal{L}$-formulas are constrained by the *formal* rules that generate them, so they can never perfectly replicate the abundant intension of mathematical statements. This is an inevitable shortcoming of our attempt to study mathematical logic and statements via a formal language.

Moreover, the Completeness Theorem is important in not only its consequences, but also its proof. Specifically, the proof of Lemma 3.17 (and thereby Theorem 3.15) provides an example of structure construction, which is crucial in proving many theorems that claim the existence of certain structures. The following theorems, which are alternative versions of Theorem 2.29 and Theorem 2.30 respectively, are typical examples:

Theorem 3.22. *(Downward Löwenheim-Skolem Theorem) Every consistent set of $\mathcal{L}_{\mathcal{A}}$-sentences has a model of size at most $||\mathcal{L}_{\mathcal{A}}||$.*

Proof. In the proof of Theorem 3.15, we may choose $\mathcal{M}^*$ so that every element of M^* is a constant, and we have $|M| = |M^*| \leq ||\mathcal{L}_{\mathcal{B}}|| = ||\mathcal{L}_{\mathcal{A}}||$. □

Theorem 3.23. *(Upward Löwenheim-Skolem Theorem) Let Γ be a set of $\mathcal{L}_{\mathcal{A}}$-sentences. If Γ has infinite models, then it has infinite models of any given power $\alpha \geq ||\mathcal{L}_{\mathcal{A}}||$.*

Proof. Let $\{c_i\}, i < \alpha$ be a sequence of distinct constant symbols not in $\mathcal{L}_{\mathcal{A}}$, and we define the set of sentences

$$\Sigma = \Gamma \cup \{(\neg(c_m \equiv c_n)) : m < n < \alpha\}.$$

Every finite subset Σ' of Σ involves finite number of constants in the sequence $\{c_i\}$. Hence any infinite model of Γ can be expanded to a model of Γ'. By the Compactness Theorem, Σ has a model $\mathcal{M} = (M, I)$, and by Theorem 3.22, this model is of size at most

$$|\mathcal{A} \cup \{c_i : i < \alpha\}| = \alpha.$$

On the other hand, the interpretations of the constants c_i in $\mathcal{M}$ must give distinct elements of M. Therefore, $\alpha \leq |M| \leq \alpha$, which means that $|M| = \alpha$. □

4. First Order Theory

We have established a deduction system for $\mathcal{L}$ and proved its soundness and completeness, thereby showing that it is an ideal system to represent mathematical deductions. However, such system is still insufficient for our inquiry. Lemma 3.9 and Theorem 3.10 show that all axioms and theorems of this system are universal truths, while many mathematical statements hold only in particular contexts, such as commutativity, expressed by $(\forall x_1(\forall x_2 F(x_1, x_2) \equiv F(x_2, x_1)))$. We want such statements to be deducible when they are true, and not deducible when they are not true, but clearly, this is beyond the capacity of our current system.

In fact, one should expect this problem once she realizes that the system defined in Definition 3.1 is a *context-free logical system*: all its axioms are logical truths, which, as we saw in Section 2, are without any mathematical meanings and out of any mathematical contexts, so from them we can deduce logical truths only. To turn it into a *mathematical system*, we need to contexualize it.

At this point, it might be tempted, as how we contextualized $\mathcal{L}$ in Section 2, to place the system under a specific mathematical structure, that is, to prescribe the universe and give each non-logical symbol an interpretation. However, this is neither necessary nor desirable. The aim of the system is to *formally* deduce one formula from another. Therefore, semantical interpretation is of no use to the system's function, unable to allow it to deduce more formulas, and could even conceal its formal nature. We thus seem to run into a paradox, where we need to make the system mathematically meaningful without assigning meaning to its formulas.

The solution is to add to the set of axioms $\mathcal{L}$-formulas that characterize certian mathematical contexts, which is illustrated by the following definition and examples.

Definition 4.1. A **first order theory** T is a formal deduction system consisting of the following four elements:

(a) a first order language $\mathcal{L}_\mathcal{A}$;
(b) the set of logical axioms and theorems within $\mathcal{L}_\mathcal{A}$;
(c) a set of $\mathcal{L}_\mathcal{A}$-sentences, which are called **non-logical axioms**;
(d) the set of rules of deduction within $\mathcal{L}_\mathcal{A}$.

Before giving the examples, the following remark should be made for clarity.

Remark 4.2. Although Definition 4.1 is a general definition of first order theory, in this paper, we treat a first order theory merely as an expansion of the first order system defined in Definition 3.1. That is, given an $\mathcal{L}_\mathcal{A}$-theory, its set of logical axioms is $\mathcal{L}_\mathcal{A} \cap \mathbb{L}$, and Theorem 4.2 is the rule of deduction. As a result, a first order theory is characterized only by its language and non-logical axioms; that is, two theories of the same language are different iff their respective sets of non-logical axioms are different.

Example 4.3. Theory of groups.

(a) The set of non-logical symbols $\mathcal{A} = \{c_0, F_\times\}$, where c_0 is a constant and $F_\times$ is a 2-variable function.

(b) The non-logical axioms are the following formulas universally quantified:

(1) $(F_\times(x_1, F_\times(x_2, x_3)) \equiv F_\times(F_\times(x_1, x_2), x_3))$;
(2) $(F_\times(x_1, c_0) \equiv x_1)$;
(3) $(F_\times(c_0, x_1) \equiv x_1)$;
(4) $(\exists x_2((F_\times(x_1, x_2) \equiv c_0) \wedge (F_\times(x_2, x_1) \equiv c_0)))$.

Example 4.4. Theory of elementary arithmetics.

(a) The set of non-logical symbols $\mathcal{A} = \{c_0, F_s, F_+, F_\times, P_<\}$, where c_0 is a constant, F_s is a 1-variable function, $F_+, F_\times$ are 2-variable functions, and $P_<$ is a 2-variable relation.

(b) The non-logical axioms are the following formulas universally quantified:

(1) $(\neg(F_s(x_1) \equiv c_0))$;
(2) $((\neg(x_1 \equiv c_0)) \to (\neg(\forall x_2(\neg(x_1 \equiv F_s(x_2))))))$;
(3) $((F_s(x_1) \equiv F_s(x_2)) \to (x_1 \equiv x_2))$;
(4) $(F_+(x_1, c_0) \equiv x_1)$;
(5) $(F_+(x_1, F_s(x_2)) \equiv F_s(F_+(x_1, x_2)))$;
(6) $(F_\times(x_1, c_0) \equiv c_0)$;
(7) $(F_\times(x_1, F_s(x_2)) \equiv F_+(F_\times(x_1, x_2), x_1))$;
(8) $(\neg(P_<(x_1, c_0)))$;
(9) $(P_<(x_1, F_s(x_2)) \to ((\neg(x_1 \equiv x_2)) \to P_<(x_1, x_2)))$;

(10) $(((\neg(x_1 \equiv x_2)) \to P_<(x_1, x_2)) \to P_<(x_1, F_s(x_2)))$;
(11) $(P_<(x_1, x_2) \to P_<(F_s(x_1), F_s(x_2)))$;
(12) $(P_<(F_s(x_1), F_s(x_2)) \to P_<(x_1, x_2))$;
(13) $((\neg(x_1 \equiv x_2)) \to ((\neg P_<(x_2, x_1)) \to P_<(x_1, x_2)))$.

Indeed, the non-logical axioms themselves do not contain mathematical meaning. However, they become meaningful once we realize that they represent mathematical statements that characterize certain mathematical contexts. In Example 4.3, the 4 axioms added correspond to the defining properties of a group, and the theorems deducible from them, by the Completeness Theorem, must also be true for any group. That is, these non-logical axioms, like a structure, create a context for the deduction system so that the system is able to deduce formulas that are true *in the context*. However, unlike a structure, which contextualizes $\mathcal{L}$ *semantically* by assigning meanings to symbols, theory does so *formally* by adding non-logical axioms. Structures and theories represent mathematical contexts in fundamentally different ways.

More importantly, we should note that a theory has limited deduction power, as the set of its theorems has been prescribed by its axioms and rules of deduction. Then, by the Completeness Theorem, the set of truths in the context that are deducible is fixed. However, such set is not necessarily equal to the set of all truths within the context; that is, the theory might be unable to deduce all truths about the context, which means that its representation of the context might be imperfect.

Therefore, it would be natural to ask, given a theory, are we able to deduce all truths within the context it specifies, or, equivalently, is there a formula true in the context but not deducible? Note that given a sentence φ and a theory T, if neither φ nor $(\neg\varphi)$ is a theorem of T, then by the Completeness Theorem, from the truth of T we are unable to determine the truth of φ and vice versa. Therefore, an answer to the above question could also uncover the logical relations between mathematical statements. We start our inquiry from the following definition:

Definition 4.5. A first order theory T is **complete** iff for any $\mathcal{L}_\mathcal{A}$-sentence θ,

$$T \vdash \theta \iff T \nvdash (\neg\theta).$$

Remark 4.6. Note that the completeness in the Completeness Theorem and the completeness in the above definition are not the same notion. The former unifies satisfiability and deducibility in the first order system, while the latter only concerns the deducibility of first order theories.

Therefore, a theory T is complete iff T is consistent and for any sentence θ, either $T \vdash \theta$ or $T \vdash (\neg\theta)$, that is, iff

$$\{\theta | \theta \text{ is a sentence, and } T \vdash \theta\}$$

is maximal consistent. Then if T is not complete, there exists a sentence φ such that either 1) $T \vdash \varphi$ and $T \vdash (\neg\varphi)$ (that is, T is not consistent) or 2) $T \nvdash \varphi$ and $T \nvdash (\neg\varphi)$. The latter possibility leads to the following definition:

Definition 4.7. Let T be a consistent first order theory, and θ be an $\mathcal{L}_\mathcal{A}$-sentence. θ is **independent** from T iff $T \cup \{\theta\}$ and $T \cup \{(\neg\theta)\}$ are consistent.

By the two definitions above, we can reformulate our question as: which theories are complete and which theories are not, and which statements are independent from an incomplete theory? The following theorem provides a partial (and thus unsatisfiable[12]) answer to these questions.

Theorem 4.8. *Let T be a consistent $\mathcal{L}_{\mathcal{A}}$-theory.*

(1) An $\mathcal{L}_{\mathcal{A}}$-sentence θ is independent from T iff there exist two models $\mathcal{M}_0, \mathcal{M}_1$ of T such that

$$\mathcal{M}_0 \vDash T \cup \{\theta\} \text{ and } \mathcal{M}_1 \vDash T \cup \{(\neg\theta)\}.$$

(2) T is complete iff if $\mathcal{M} \vDash T$ and θ is an $\mathcal{L}_{\mathcal{A}}$-sentence, then

$$T \vdash \theta \iff \mathcal{M} \vDash \theta.$$

(3) T is complete iff any two models of T are elementary equivalent.

(4) T is complete iff there exists a model $\mathcal{M}$ of T such that if θ is an $\mathcal{L}_{\mathcal{A}}$-sentence, then

$$T \vdash \theta \iff \mathcal{M} \vDash \theta.$$

Proof. (1) ($\Rightarrow$) Suppose θ is independent from T; that is, $T \nvdash \theta$ and $T \nvdash (\neg\theta)$. By Theorem 3.8, $T \cup \{\theta\}$ and $T \cup \{(\neg\theta)\}$ are both consistent. By the Completeness Theorem, they both have a model.

($\Leftarrow$) Suppose $\mathcal{M}_0, \mathcal{M}_1$ are models of T, and they are models of θ and $(\neg\theta)$ respectively. Suppose $T \vdash \theta$. By the Completeness Theorem, $T \vDash \theta$, so $\mathcal{M}_1 \vDash \theta$. But $\mathcal{M}_1 \vDash (\neg\theta)$, which is a contradiction. Thus, $T \nvdash \theta$. For the same reason, $T \nvdash (\neg\theta)$.

(2) ($\Rightarrow$) Suppose T is complete, $\mathcal{M} \vDash T$. Let θ be an $\mathcal{L}_{\mathcal{A}}$-sentence. If $T \vdash \theta$, by the Completeness Theorem, $T \vDash \theta$, so $\mathcal{M} \vDash \theta$. If $T \nvdash \theta$, then $T \vdash (\neg\theta)$. By the Completeness Theorem, $T \vDash (\neg\theta)$, so $\mathcal{M} \vDash (\neg\theta)$ and $\mathcal{M} \nvDash \theta$.

($\Leftarrow$) Suppose T is incomplete. Let θ be a sentence independent from T. Let $\mathcal{M}_0 \vDash T \cup \{\theta\}, \mathcal{M}_1 \vDash T \cup \{(\neg\theta)\}$.

(3) ($\Rightarrow$) Suppose T is complete, and $\mathcal{M}_0 \vDash T, \mathcal{M}_1 \vDash T$. Let θ be an $\mathcal{L}$-sentence. By (2),

$$\mathcal{M}_0 \vDash \theta \iff T \vdash \theta \iff \mathcal{M}_1 \vDash \theta.$$

($\Leftarrow$) Suppose T is not complete. Let θ be a sentence independent from T. By (1), let $\mathcal{M}_0 \vDash T \cup \{\theta\}, \mathcal{M}_1 \vDash T \cup \{(\neg\theta)\}$. Then $\mathcal{M}_0$ and $\mathcal{M}_1$ are not elementary equivalent.

(4) The proof is almost the same as (2). □

Interestingly, we saw previously that theories and structures are fundamentally different representations of mathematical contexts. However, the theorem above reveals that they are in fact related to each other. Recall that a theory prescribes a set of sentences that it is able to deduce, and thus, by the Completeness Theorem, a set of truths within the mathematical context it represents; meanwhile, a structure also prescribes a set of truths within the context it represents. Therefore, they are related as there could be certain relations between the sets of truths they prescribe respectively. And Theorem 4.8(4) implies that when these two sets coincide, that is, when the mathematical contexts represented by the theory and the structure become indistinguishable in $\mathcal{L}$, the theory is complete.

[12]See Conclusion.

5. Conclusion

I have introduced some fundamental concepts and results in first order logic and model theory. From their underlying rationale and relations with each other we are able to spy on the intriguing world of logic, mathematics, and their philosophy.

However, due the introductory nature of this paper and the length limit, I was unable to dive too deeply into details about the topics I mentioned. I omitted the proofs of certain results, such as $(\mathbb{Q}, <) \preceq (\mathbb{R}, <)$, as well as definitions and theorems that could arguably be more important than those I have introduced. Moreover, in Section 4, I only introduced very limited number of concepts and did not led the discussion very far. However, in my opinion, what characterizes a good introduction to mathematical logic is not a wide coverage of topics, but a detailed discussion of the motivations and implications behind the definitions and theorems, which has been the main focus of this paper. Indeed, mathematical logic is an abstraction of mathematics, while mathematics is abstract by itself; thus, to understanding the former, it is crucial to understand the concrete motivations and intuition underlying various constructions and abstractions.

Nevertheless, to remedy the limited coverage of this paper, I will point at some directions to which it leads. Indeed, the problem in section 4 has not attained a satisfying solution. We need means to prove whether a theory is complete and whether a sentence is independent, while what Theorem 4.8 offers is very limited: by the theorem, to prove completeness or independence we need to examine the models of a theory; however, the set of structures is so large that it would be difficult to assemble all models of any given theory. Therefore, to solve the problem, we need to look for other properties and conditions, and this is where **quantifier elimination** and **minimal substructure** become useful.

Moreover, the implication of the Compactness Theorem is so deep that what could be uncovered are much more than Theorem 3.21. Examples include the existence of non-standard model of complete number theory and $\mathcal{L}$'s inability to delineate well-ordering.

Finally, I invite my readers to think about one more question, which, I believe, is of no less significance than any problem brought up previously. We study mathematical logic as a field of mathematics, in which we rigorously define concepts and prove theorems. However, such approach could be fundamentally problematic because of its circularity. Specifically, to prove truths about mathematical logic, we use mathematical logic itself, so the truths we thus attain might not be valid as we have already assumed them in our proofs! This problem seems unavoidable for any mathematical approach to studying mathematical logic. But is there a better approach?

Acknowledgments. Planning, drafting, and revising this paper took a significant amount of time and effort, but if there were only myself, the work could not have been finished.

I want to thank the organizers of the 2021 REU, which provided me an opportunity to study mathematical logic, ponder it, and write a serious paper about it. I also want to thank Mr. Chang, Mr. Keisler, and Mr. Feng for the excellent textbooks. Their elegant proofs and detailed explanations have made them great sources for reference and inspiration.

Special gratitude is given to my mentor Connor Lockhart and professor Peter May. With Connor I had enlightening conversations about logic and its philosophy, and he also helped edit and revise my drafts with many valuable suggestions. Moreover, the first draft of this paper contained many errors and typos due to hurried submission, but professor May carefully pointed them out in his review. The subsequent revision was tiredsome and required significant amount of effort and patience, and without May's encouragement and motivation, the work could have not been done.

References

[1] C. C. Chang and H. Jerome Keisler. Model Theory. Elsevier Science Publishers B.V.. 1990.
[2] Qi Feng. Introduction to Mathematical Logic. Science Press. 2017.