First order Logic

kankanray

Contents

1 Propositional Logic

1.1 Syntax

We have logic symbols $\neg$ and $\rightarrow$, (and).

We also have a set of proposition letters $A_0, A_1, \ldots$. We often assume we have countably many proposition letters.

We consider the set of all finite sequences of proposition letters and logic symbols, and select some suitable ones from them.

Definition 1.1. A a well-formed formula(or sentence) φ are defined by:

$\varphi ::= A_n|(\neg\psi)|(\psi \rightarrow \phi)$.

Denote $\mathcal{L}_0$ as the set of all well-formed formulas.

We know $\mathcal{L}$ is the smallest set that is closed under negation and implication, and contains all proposition letters.

We have several lemmas:

Lemma 1.1. For every $\varphi \in \mathcal{L}_0$, φ is exactly one of the three cases:
(1) φ is a proposition letter A_n,
(2) There is a $\psi \in \mathcal{L}_0$, such that $\varphi = (\neg\psi)$,
(3) There are $\psi \in \mathcal{L}_0$ and $\phi \in \mathcal{L}_0$ such that φ is $(\psi \to \phi)$

This one is obvious, just o through the definition of $\mathcal{L}_0$.

The following two lemmas strengthen our understanding of formulas in $\mathcal{L}_0$:

Lemma 1.2. For sentences $\varphi \in \mathcal{L}_0$, all proper prefix of φ is not a well-formed formula.

Here, φ is a sequence of symbols and letters, essentially, so proper prefix of φ is a sub-sequence, start from the head of φ, with no one is left in the middle, but shorter than φ.

This lemma can be proved by induction on complexity of formulas.

Then we obtain a useful theorem:

Theorem 1.1. For every $\varphi \in \mathcal{L}_0$, φ is exactly one of the three cases:
(1) φ is a unique proposition letter A_n,
(2) There is a unique $\psi \in \mathcal{L}_0$, such that $\varphi = (\neg\psi)$,
(3) There are unique $\psi \in \mathcal{L}_0$ and $\phi \in \mathcal{L}_0$ such that φ is $(\psi \to \phi)$

1.2 Semantics

After given the syntax of propositional logic, we need semantics.

We start from the valuation on proposition letters:

Definition 1.2. A *valuation* ν is a map from the set of proposition letters to the set $\{0, 1\}$. (1 means true, and 0 means false)

To extend the valuation to all formulas in $\mathcal{L}_0$, we need the following 2 functions:

Definition 1.3. $H_\neg : \{0, 1\} \to \{0, 1\}$, with $H_\neg(0) = 1$ and $H_\neg(1) = 0$.

$H_\to : \{0, 1\} \times \{0, 1\} \to \{0, 1\}$, with $H_\to(0, 0) = H_\to(0, 1) = H_\to(1, 1) = 1$, and $H_\to(1, 0) = 0$.

Then we have an obvious theorem:

Theorem 1.2. Given valuation ν on proposition letters, there is a unique extension $\bar{\nu}$ on $\mathcal{L}_0$, satisfying:
$\bar{\nu}(A_n) = \nu(A_n)$, for all natural number n,
$\bar{\nu}((\neg\psi)) = H_\neg(\bar{\nu}(\psi))$, and
$\bar{\nu}((\psi \to \phi)) = H_\to(\bar{\nu}(\psi), \bar{\nu}(\phi))$, for all formulas ψ and ϕ.

For some specific formulas, we need not know valuations on all proposition letters. We need only the letters that appear in those formulas. We have a theorem describing this:

Theorem 1.3. Given $\varphi \in \mathcal{L}_0$, and two valuation function ν and μ. Suppose for all proposition letters A_n that appears in φ, we have $\nu(A_n) = \mu(A_n)$. Then $\bar{\nu}(\varphi) = \bar{\mu}(\varphi)$.

We denote the relation $\Gamma \models \varphi$.

This can be by induction easily.

More corollaries on valuation and substitution are omitted, since they are quite obvious.

We have some basic concepts:

Definition 1.4. (1) A valuation ν *satisfies* a formula φ iff $\bar{\nu}(\varphi) = 1$.

(2) A valuation ν *satisfies* a set of formulas Γ iff for all formulas φ in Γ, we have $\bar{\nu}(\varphi) = 1$.

(3) A formula φ is said to be *satisfiable* iff there is a valuation satisfies it.

(4) A set of formulas Γ is said to be *satisfiable* iff there is a valuation satisfies it.

(5) A formula is called a *tautology* iff all valuation satisfies it.

(6) A formula is called a *contradiction* iff no valuation satisfies it.

Definition 1.5. Given a set of formulas Γ, and a formula φ, we say φ is a logical consequence iff for every valuation ν, if ν satisfies Γ, then ν satisfies ν also.

Definition 1.6. Two formulas φ and ψ are said to be equivalent iff $\{\varphi\} \models \psi$ and $\{\psi\} \models \varphi$.

Besides $\neg$ and $\to$, we often use *conjunction* $\wedge$ and *disjunction* $\vee$, and *equality* $\leftrightarrow$. They can all be defined with $\neg$ and $\to$:

Definition 1.7. $(\psi \wedge \phi) := \neg((\psi \to (\neg\phi))$,

$(\psi \vee \phi) := ((\neg\psi) \to \phi)$,

$(\psi \leftrightarrow \phi) : ((\psi \to \phi) \wedge (\phi \to \psi))$.

There valuation are quite obvious, thus omitted.

1.3 Deduction System

In this section we discuss deduction system of propositional logic, and in the next section, will show the soundness and completeness of it.

Definition 1.8. Logical axioms

Assume A, B, C are formulas in $\mathcal{L}_0$

(I)

1. $(A \to A)$,

2. $(((\neg A) \to A) \to A)$,
(II)
1. $(A \to (B \to A))$,
2. $(A \to ((\neg A) \to B))$,
3. $((\neg A) \to (A \to B))$,
4. $(A \to ((\neg B) \to (\neg(A \to B))))$,
(III)
1. $((A \to (B \to C)) \to ((A \to B) \to (A \to C)))$.

We can easily verify that all logical axioms are tautologies.
Now we consider formal proofs and theorems.

Definition 1.9. Given a set of formulas Γ, and a sequence of formulas $\langle B_1, ..., B_n\rangle$. We say such sequence is a proof, iff for all $1 \leq i \leq n$,
(a) B_i is a logical axiom, or a formula in Γ.
(b) There are $j, k < i$, such that B_k is just $(B_j \to B_i)$.

Definition 1.10. A formula φ is said to be a theorem of Γ iff there is a proof $\langle B_1, ..., B_n\rangle$ of Γ, with $B_n = \varphi$.
We use $\Gamma \vdash \varphi$ to describe this.

We prove a simple fact:

Proposition 1.1.
$\vdash ((\neg(\neg\varphi)) \to \varphi)$

Proof. Omitted. □

Then we have three useful lemmas, especially the *Deduction Lemma*.

Lemma 1.3.
If $\Gamma \vdash A$ and $\Gamma \vdash (A \to B)$, then $\Gamma \vdash B$.

Lemma 1.4. (Deduction Lemma)
$\Gamma \cup \{\varphi\} \vdash \psi$ iff $\Gamma \vdash (\varphi \to \psi)$

Proof. One side is direct. The other side requires induction. □

Lemma 1.5.
If $\Gamma \cup \{\neg\varphi\} \vdash \varphi$, then $\Gamma \vdash \varphi$

1.4 Completeness

We start from the concept of *consistency*.

Definition 1.11. A set of formulas Γ is *inconsistent* if there is a formula φ such that
$\Gamma \vdash \varphi$ and $\Gamma \vdash (\neg\varphi)$.

Actually every formula is a theorem of an inconsistent set of formulas Γ:

Lemma 1.6. If Γ is inconsistent, then $\Gamma \vdash \varphi$ for every $\varphi \in \mathcal{L}_0$.

Definition 1.12. $\Gamma \subseteq \mathcal{L}_0$ is *consistent* if it's not inconsistent.

Definition 1.13. Given consistent set $\Gamma \subseteq \mathcal{L}_0$ and a formula φ, we say φ is independent from Γ iff $\Gamma \nvdash \varphi$ and $\Gamma \nvdash (\neg\varphi)$

We have an useful compactness theorem:

Theorem 1.4. (Compactness)
$\Gamma \subseteq \mathcal{L}_0$ is consistent iff every finite subset of Γ is consistent.

Consistent sets also have some basic properties:

Lemma 1.7. Given consistent subset $\Gamma \subseteq \mathcal{L}_0$, and a formula φ, then one of the following three must hold:
(1) $\Gamma \cup \{\varphi\}$ is consistent,
(2) $\Gamma \cup \{(\neg\varphi)\}$ is consistent,
(3) $\Gamma \cup \{\varphi\}$ and $\Gamma \cup \{(\neg\varphi)\}$ are consistent.

Lemma 1.8. If $\Gamma \subset \mathcal{L}_0$ is consistent, and φ is independent to φ, then $\Gamma \cup \{\varphi\}$ and $\Gamma \cup \{(\neg\varphi)\}$ are consistent.

To prove completeness, we need a special kind of consistent set, called the maximally consistent set:

Definition 1.14. We say a consistent set $\Gamma \subseteq \mathcal{L}_0$ is *maximally consistent*, iff for every formula φ, if $\Gamma \cup \{\varphi\}$ is consistent, then $\varphi \in \Gamma$.

Maximally consistent set has good properties. Actually, such set is satisfiable:

Lemma 1.9. If Γ is maximally consistent, then it is satisfiable.

Proof. To prove Γ is satisfiable, we need to construct a valuation function. Simply consider ν defined as the following:

$$\nu(A_n) = \begin{cases} 1 & \text{if } A_n \in \Gamma \\ 0 & \text{if } A_n \notin \Gamma \end{cases} \tag{1}$$

Then we can prove by induction that $\varphi \in \Gamma$ iff $\bar{\nu}(\varphi) = 1$, for every formula $\varphi \in \mathcal{L}_0$. so ν satisfies Γ. $\square$

What about consistent set in general? The next lemma says that every consistent set Γ can be extended to a maximally consistent set Γ^*, with $\Gamma \subseteq \Gamma^*$:

Lemma 1.10. Given $\Gamma \subseteq \mathcal{L}_0$ consistent. Then there exists a maximally consistent $\Gamma^* \subseteq \mathcal{L}_0$, with $\Gamma \subseteq \Gamma^*$.

Proof. Since we have only countably infinite proposition letters, with finite logical symbols, $\mathcal{L}_0$ is a subset of all finite sequences of proposition letters and logical symbols, so $\mathcal{L}_0$ is countably infinite. Then we list all the formulas in $\mathcal{L}_0$: $\varphi_0, \varphi_1, ..., \varphi_n, ...$.

We start from $n = 0$. Let $\Gamma_0 = \Gamma$. If $\Gamma_0 \cup \{\varphi_0\}$ is consistent, then $\Gamma_1 := \Gamma_0 \cup \{\varphi_0\}$. Otherwise, let $\Gamma_1 = \Gamma_0$.

Assume we already have Γ_n, then if $\Gamma_n \cup \{\varphi_n\}$ is consistent, then $\Gamma_{n+1} := \Gamma_n \cup \{\varphi_n\}$. Otherwise, let $\Gamma_{n+1} = \Gamma_n$.

We define $\Gamma^* = \bigcup\{\Gamma_n | n \in \mathbb{N}\}$. And it's not hard to show that Γ^* is maximally consistent, with $\Gamma \subseteq \Gamma^*$. □

So we have the satisfiable condition for arbitrary consistent set:

Theorem 1.5. If $\Gamma \subseteq \mathcal{L}_0$ is consistent, then it's satisfiable.

Proof. We have $\Gamma^* \subseteq \mathcal{L}_0$ maximally consistent with $\Gamma \subseteq \Gamma^*$, thus satisfied by ν. Then it's clear that ν satisfies Γ also. □

Another question is, if Γ is satisfiable, is it consistent? To answer such question, we have need *soundness* theorem:

Theorem 1.6. If $\Gamma \vdash \varphi$, then $\Gamma \models \varphi$.

Proof. Use induction on the length of proof. $\Gamma \vdash \varphi$ means that there is a proof $\langle B_1, ..., B_n \rangle$, with $B_n = \varphi$.

Base step: $B_1 \in \Gamma$ or is a logical axiom, so $\Gamma \models B_1$.

Induction step: assume $\Gamma \models B_i$ for $1 \leq i \leq k < (n)$, then if $B_{k+1} \in \Gamma$ or is a logical axiom, then $\Gamma \models B_{k+1}$ also. If there is $i, j \leq (k+1)$, with $B_j = (B_i \to B_{k+1})$, then $\Gamma \models B_i$ and $\Gamma \models (B_i \to B_{k+1})$, so we have $\Gamma \models B_{k+1}$.

By induction, $\Gamma \models B_n$, i.e. $\Gamma \models \varphi$. □

Now we can answer questions above:

Theorem 1.7. If $\Gamma \subseteq \mathcal{L}_0$ is satisfiable, then it's consistent.

Proof. Otherwise, assume Γ is inconsistent, then there exists φ such that $\Gamma \vdash \varphi$ and $\Gamma \vdash (\neg\varphi)$, then by soundness theorem, $\Gamma \models \varphi$ and $\Gamma \models (\neg\varphi)$.

Since Γ is satisfiable, there is a valuation ν satisfies Γ, so it satisfies φ and $(\neg\varphi)$ also, which is impossible.

So Γ is consistent. □

As a corollary, we have another version of compactness property:

Corollary 1.1. $\Gamma \subseteq \mathcal{L}_0$ is satisfiable iff all its finite subset is satisfiable.

Finally, we have *completeness* theorem for propositional logic:

Theorem 1.8. If $\Gamma \models \varphi$, then $\Gamma \vdash \varphi$.

Proof. Assume $\Gamma \not\models \varphi$, then by a lemma introduced above, we must have $\Gamma \cup \{(\neg\varphi)\}$ is consistent: otherwise $\Gamma \cup \{(\neg\varphi)\} \vdash \varphi$, then $\Gamma \vdash \varphi$, contradiction.

So there is a valuation ν satisfies $\Gamma \cup \{(\neg\varphi)\}$, so we have $\bar{\nu}(\neg\varphi) = 1$ and ν satisfies Γ. Since $\Gamma \models \varphi$, we have $\bar{\nu}(\varphi) = 1$ also, which is a contradiction.

Therefore we must have $\Gamma \vdash \varphi$. $\square$

Thus, the propositional logic ends here.

2 First order Logic Basic

2.1 Syntax

In first order logic, we have more symbols than in propositional logic.

-Logical symbols: $\neg$, $\to$, (,), and $\forall$.

-Equation symbol: $\hat{=}$.

-Variable symbols: for every natural number $n \in \mathbb{N}$, we have a variable symbol x_n.

-Constant element symbols: for every natural number $n \in \mathbb{N}$, we have a constant symbol c_n.

-Function symbols: for every natural number $n \in \mathbb{N}$, we have a function symbol F_n.

-Predicate symbol: for every natural number $n \in \mathbb{N}$, we have a predicate symbol P_n.

We assume those symbols are pair-wisely different. Also, we suppose that for every natural number n, there is unique positive integer for F_n, say $\pi(F_n)$, representing its 'dimension'. For every positive integer k, assume that there infinitely many function symbols with dimension k. When $\pi(F_n) = k$, we say F_n is a k-ary function symbol. We have similar assumptions on predicate symbols also($\pi(P_n)$ and so on).

With those various symbols, we can construct finite symbol sequences, and pick out the meaningful ones from among them.

Note that we can formally define what finite symbol sequence is, what is a prefix, what is a proper prefix, what is the combination of two finite sequences(A's tail is connected to B's head). But that's boring. So we'll introduce our language sloppily. I believe that readers know what I mean.

We first use variable and constant element symbols, together with function symbols to define *terms*:

Definition 2.1. (Terms)

-x_n, c_n are terms.

-If $\tau_1, ..., \tau_k$ are terms, and F_n is a k-ary function symbol, then $F_n(\tau_1, ..., \tau_n)$ is a term.

-Terms can only be constructed by the above two methods.

We denote the set of all terms as T.

As in propositional logic, we have similar results on *readability*:

Lemma 2.1. If τ is a term, then it must be one of the two cases:

(1) τ is some x_n or c_n.

(2) There are terms $\tau_1, ..., \tau_k$, and a k-ary function symbol F_n such that τ is $F_n(\tau_1, ..., \tau_n)$.

Lemma 2.2. If τ is a term, then any proper prefix it NOT a term.

Theorem 2.1. If τ is a term, then it must be one of the two cases:

(1) τ is some unique x_n or c_n.

(2) There are unique terms $\tau_1, ..., \tau_k$, and a unique k-ary function symbol F_n such that τ is $F_n(\tau_1, ..., \tau_n)$.

After defining what a term is, we now define formulas in first order logic.

Definition 2.2. (Formulas)

-If τ_1 and τ_2 are terms, then $(\tau_1 \hat{=} \tau_2)$ is a formula.

-If P_n is a k-ary predicate symbol, and $\tau_1, ..., \tau_k$ are terms, then $P_n(\tau_1, ..., \tau_k)$ is a formula.

-If φ is a formula, then $(\neg\varphi)$ is a formula.

-If ψ and ϕ are formulas, then $(\psi \to \phi)$ is a formula.

-If φ is a formula, x_n is a variable symbol, then $(\forall x_n \varphi)$ is a formula.

-Formulas can only be constructed by the above methods.

We denote the set of all formulas as $\mathcal{L}$.

We also have some readability properties:

Lemma 2.3. If φ is a formula, then it must one of the following cases:

(1) There are terms τ_1 and τ_2, such that φ is $(\tau_1 \hat{=} \tau_2)$.

(2) There is a k-ary predicate symbol P_n, and terms $\tau_1, ..., \tau_k$ such that φ is $P_n(\tau_1, ..., \tau_k)$.

(3) There is a formula ψ such that φ is $(\neg\psi)$.

(4) There are formulas ψ and ϕ such that φ is $(\psi \to \phi)$.

(5) There is a formula ψ and a variable symbol x_n such that φ is $(\forall x_n \psi)$

Lemma 2.4. If φ is a formula, then no proper prefix of φ is a formula.

Theorem 2.2. If φ is a formula, then it must one of the following cases:

(1) There are unique terms τ_1 and τ_2, such that φ is $(\tau_1 \hat{=} \tau_2)$.

(2) There is a unique k-ary predicate symbol P_n, and unique terms $\tau_1, ..., \tau_k$ such that φ is $P_n(\tau_1, ..., \tau_k)$.

(3) There is a unique formula ψ such that φ is $(\neg\psi)$.

(4) There are unique formulas ψ and ϕ such that φ is $(\psi \to \phi)$.

(5) There is a unique formula ψ and a unique variable symbol x_n such that φ is $(\forall x_n \psi)$

An important concept here is *substitution*.

Definition 2.3. Given a formula φ and a symbol a, an *occurrence* of a in φ is defined in the obvious way.

We say $\forall x_n$ occurs iff $\forall$ occurs at j and x_n occurs at $j + 1$.

Definition 2.4. Range of '$\forall$'

Given a formula φ, where $\forall x_n$ occurs somewhere in it. Then the range of $\forall x_n$ in the occurrence is, roughly speaking, the *range* is from (to) in ...($\forall x_n$...)....

To be more specific, assume (appears at $j-1$, $\forall$ appears at j, and x_n appears at $j+1$, and consider the) pairing to (at $j-1$, then the range of $\forall x_n$ is said to be from $j-1$ to k

Definition 2.5. Free and bounded

1. A free occurrence of x_n is an occurrence of x_n, and this x_n is not in any range of $\forall x_n$.

2. A bounded occurrence of x_n is an occurrence of x_n that is not a free occurrence.

3. A variable x_n is said to be a free variable of φ iff x_n has a free occurrence in φ.

4. A variable x_n is said to be a bounded variable of φ iff x_n occurs in φ and it's NOT a free variable.

We use $\tau(x_1,...,x_n)$ to express that all variables occur in τ are contained in $\{x_1,...,x_n\}$

We use $\varphi(x_1,...,x_n)$ to express that all free variables occur in φ are contained in $\{x_1,...,x_n\}$

Definition 2.6. A formula is said to be a *sentence* if it has no free variables.

Definition 2.7. Universalization

Given $\varphi(x_1,...,x_n)$, we get a sentence $(\forall x_{i_1}...(\forall x_{i_n}\varphi(x_1,...,x_n)))$ is a universalization, where $\{x_{i_1},...,x_{i_n}\} = \{x_1,...,x_n\}$.

Definition 2.8. Substitution

Assume $x_1,...,x_n$ are variables, $\tau_1,...,\tau_n$ are terms, $\tau(x_1,...,x_n)$ is a term, and $\varphi(x_1,...,x_n)$ is a formula.

1. We use $\tau(x_1,...,x_n;\tau_1,...,\tau_n)$ to express we simultaneously substitute all occurrences of $x_1,...,x_n$ by $\tau_1,...,\tau_n$.

2. We use $\varphi(x_1,...,x_n;\tau_1,...,\tau_n)$ to express we simultaneously substitute all free occurrence of $x_1,...,x_n$ by $\tau_1,...,\tau_n$.

It's clear that after substitution, we still have a term (or a formula).

Lemma 2.5. Assume $x_1,...,x_n$ are variables, $\tau_1,...,\tau_n$ are terms, $\tau(x_1,...,x_n)$ is a term, and $\varphi(x_1,...,x_n)$ is a formula.

1. $\tau(x_1,...,x_n;\tau_1,...,\tau_n)$ is a term.
2. $\varphi(x_1,...,x_n;\tau_1,...,\tau_n)$ is a formula.

When do substitution in terms, nothing changes maybe. While handling with formulas, some substitution may totally change the 'meaning' of formulas before substitution. So must consider proper substitution only:

Definition 2.9. Substitutablity

Suppose φ is formula, x_n is a free variable of φ, and τ is a term, we say τ can substitute x_n in φ iff for every variable x_j that occurs in τ, if $\forall x_j$ occurs in φ also, then there is no free occurrence of x_n in any range of $\forall x_j$ in φ.

Note that $\wedge$, $\vee$, $\leftrightarrow$ are defined as in propositional logic. We can also define $\exists$ as:

$\exists x_n \varphi := (\neg(\forall x_n (\neg\varphi)))$

2.2 Semantics

Let $\mathcal{A}$ be a set containing some constant element symbols, function symbols and predicate symbols. We denote the set of all terms in T whose constant element symbols and function symbols are in $\mathcal{A}$ as $T_{\mathcal{A}}$, and the set of all formulas in $\mathcal{L}$ whose terms are from $T_{\mathcal{A}}$ and predicate symbols are from $\mathcal{A}$, as $\mathcal{L}_{\mathcal{A}}$

Definition 2.10. A *structure* of $\mathcal{L}_{\mathcal{A}}$ is a binary tuple (M, I), where
1. M is an non-empty set(we call it *domain*).
2. $I : \mathcal{A} \to M \cup \bigcup\{\mathfrak{P}(M^n)|n \in \mathbb{N}\}$(we call it *interpretation*) satisfying:
-If $c_n \in \mathcal{A}$, then $I(c_n) \in M$,
-If $F_n \in \mathcal{A}$ and $\pi(F_n) = k$, then $I(F_n) : M^K \to M$, and
-If $P_n \in \mathcal{A}$ and $\pi(P_n) = k$, then $I(P_n) \subseteq M^k$.

In other words, $I(c_n)$ is an element of M, $I(F_n)$ is a k-ary function, and $I(P_n)$ is a k-ary relation.

To tell a formula is right or wrong, we need some 'valuation' on variable symbols.

Let $\mathcal{B}$ be the set of all variable symbols.

Definition 2.11. Given $\mathcal{M} = (M, I)$ a structure of $\mathcal{L}_{\mathcal{A}}$, we call a map $\nu : \mathcal{B} \to M$ a *valuation* function.

Note that we can extent a valuation function ν on variable symbols to a valuation $\bar{\nu}$ on all terms, recursively:
-$\bar{\nu} = \nu(x_n)$,
-$\bar{\nu}(c_n) = I(c_n)$,
-$\bar{\nu}(F_n(\tau_1, ..., \tau_k)) = I(F_n)(\bar{\nu}(\tau_1), ..., \bar{\nu}(\tau_k))$

As in propositional logic, valuation are sometimes 'determined' locally:

Definition 2.12. Given a structure $\mathcal{M}$ of $\mathcal{L}_{\mathcal{A}}$, a term $\tau \in T_{\mathcal{A}}$, and a formula $\varphi \in \mathcal{L}_{\mathcal{A}}$, and two valuation functions ν and μ,

1. We say $\nu \equiv_\tau \mu$ or $\nu = \mu \bmod(\tau)$ iff for every x_j occurs in τ, we have $\nu(x_j) = \mu(x_j)$.

2. We say $\nu \equiv_\varphi \mu$ or $\nu = \mu \bmod(\varphi)$ iff for every x_j freely occurs in φ, we have $\nu(x_j) = \mu(x_j)$.

We then have a direct lemma:

Lemma 2.6. Given structure $\mathcal{M}$, a term $\tau \in T_{\mathcal{A}}$ and two valuation functions ν and μ. If $\nu \equiv_\tau \mu$, then
$\bar{\nu}(\tau) = \bar{\mu}(\tau)$.

Finally, we can define the semantics of first order logic:

Definition 2.13. Given a structure $\mathcal{M}$ of $\mathcal{L}_{\mathcal{A}}$, and a valuation function ν, and a formula $\varphi \in \mathcal{L}_{\mathcal{A}}$, we use $(\mathcal{M}, \nu) \models \varphi$ to express $(\mathcal{M}, \nu)$ *satisfies* φ. The details are as follows:

(1) $(\mathcal{M}, \nu) \models (\tau_1 \hat{=} \tau_2)$ iff $\bar{\nu}(\tau_1) = \bar{\nu}(\tau_2)$.
(2) $(\mathcal{M}, \nu) \models P_n(\tau_1, ..., \tau_k)$ iff $(\bar{\nu}(\tau_1), ..., \bar{\nu}(\tau_k)) \in I(P_n)$.
(3) $(\mathcal{M}, \nu) \models (\neg\varphi)$ iff $(\mathcal{M}, \nu) \not\models \varphi$.
(4) $(\mathcal{M}, \nu) \models (\psi \to \phi)$ iff $(\mathcal{M}, \nu) \not\models \psi$ or $(\mathcal{M}, \nu) \models \phi$.
(5) $(\mathcal{M}, \nu) \models (\forall x_n \varphi)$ iff for every valuation function μ, if $\nu \equiv_\varphi \mu$, then $(\mathcal{M}, \mu) \models \varphi$.

We then have a local property for $\models$ relation:

Theorem 2.3. If $\nu \equiv_\varphi \mu$, then $(\mathcal{M}, \nu) \models \varphi$ iff $(\mathcal{M}, \mu) \models \varphi$.

Recall that a formula is called a sentence if it has no free variables. Sentences has good properties on satisfaction:

Corollary 2.1. Suppose φ is a sentence, then the following two are equivalent:
(1) There is a $\mathcal{M}$-valuation ν satisfies φ.
(2) All $\mathcal{M}$-valuations satisfies φ.

By this corollary, we can talk about the following:

Definition 2.14. Given a sentence φ, a structure $\mathcal{M}$, we say:
(1) $\mathcal{M} \models \varphi$ iff there is a $\mathcal{M}$-valuation ν satisfies φ.
(2) $\mathcal{M} \not\models \varphi$ iff there is no $\mathcal{M}$-valuation ν satisfies φ.

Recall universalization, it also has good properties:

Corollary 2.2. Assume $\psi(x_1, ..., x_n)$ is a formula, and $\varphi = \forall x_1 ... \forall x_n \psi$, then the following two are equivalent:
(1) For all ν, we have $(\mathcal{M}, \nu) \models \varphi$.
(2) For all μ, we have $(\mathcal{M}, \mu) \models \psi$.

Now we discuss more about substitution:

Theorem 2.4. Given a structure $\mathcal{M} = (M, I)$ of $\mathcal{L}_{\mathcal{A}}$.
1. Given terms $\tau(x_1, ..., x_n)$, $\tau_1, ..., \tau_n$, and valuation functions μ, ν, assume
$\mu(x_i) = \bar{\nu}(\tau_i)$ for $1 \leq i \leq n$,
then we have $\bar{\mu}(\tau(x_1, ..., x_n)) = \bar{\nu}(\tau(x_1, ..., x_n; \tau_1, ..., \tau_n))$
2. Given formula $\varphi(x_1, ..., x_n)$ and terms $\tau_1, ..., \tau_n$, τ_i can substitute x_i for all $1 \leq i \leq n$, and two valuation functions μ, ν, assume
$\mu(x_i) = \bar{\nu}(\tau_i)$ for $1 \leq i \leq n$,
then we have $(\mathcal{M}, \mu) \models \varphi(x_1, ..., x_n)$ iff $(\mathcal{M}, \nu) \models \varphi(x_1, ..., x_n; \tau_1, ..., \tau_n)$.

We have more definitions:

Definition 2.15. Given a structure $\mathcal{M} = (M, I)$ of $\mathcal{L}_{\mathcal{A}}$, $b_1, ..., b_n \in M$, term $\tau(x_1, ..., x_n)$ and formula $\varphi(x_1, ..., x_n)$, we have the following definitions:

(1) We use $\tau[b_1, ..., b_n]$ to denote the unique element $\bar{\nu}(\tau)$, where $\nu(x_i) = b_i$ for $1 \leq i \leq n$.

(2) We use $\mathcal{M} \models \varphi[b_1, ..., b_n]$ to denote the fact that, for ν with $\nu(x_i) = b_1$ for $1 \leq i \leq n$, we have $(\mathcal{M}, \nu) \models \varphi$.

Note that the definition make sense, since we have local property of valuation functions.

With new definition, we can re-express substitution theorem:

Theorem 2.5. Given a structure $\mathcal{M} = (M, I)$ of $\mathcal{L}_{\mathcal{A}}$, $b_1, ..., b_n \in M$ and

1. Given terms $\tau(x_1, ..., x_n)$, $\tau_1, ..., \tau_n$, and valuation function ν, assume $b_i = \bar{\nu}(\tau_i)$ for $1 \leq i \leq n$, then we have $\tau[b_1, ..., b_n] = \bar{\nu}(\tau(x_1, ..., x_n; \tau_1, ..., \tau_n))$

2. Given formula $\varphi(x_1, ..., x_n)$ and terms $\tau_1, ..., \tau_n$, τ_i can substitute x_i for all $1 \leq i \leq n$, and valuation function ν, assume
$b_i = \bar{\nu}(\tau_i)$ for $1 \leq i \leq n$,
then we have $\mathcal{M} \models \varphi[b_1, ..., b_n]$ iff $(\mathcal{M}, \nu) \models \varphi(x_1, ..., x_n; \tau_1, ..., \tau_n)$.

Note that the proof of substitution theorem are quite long, although the idea is quite direct: use induction.

Anyway, we've finished our job in first order logic semantics.

We now introduce some concepts similar to propositional logic.

Definition 2.16. Satisfaction

Given a formula $\varphi \in \mathcal{L}_{\mathcal{A}}$, a set of formulas $\Gamma \subseteq \mathcal{L}_{\mathcal{A}}$,

(1) We say φ is satisfiable iff there exists a structure $\mathcal{M}$ of $\mathcal{L}_{\mathcal{A}}$, and a valuation function ν such that $(\mathcal{M}, \nu) \models \varphi$.

(2) We say Γ is satisfiable iff there exists a structure $\mathcal{M}$ of $\mathcal{L}_{\mathcal{A}}$, and a valuation function ν such that $(\mathcal{M}, \nu) \models \psi$ for every $\psi \in \Gamma$.

Definition 2.17. Validity

Given a structure $\mathcal{M} = (M, I)$ of $\mathcal{L}_{\mathcal{A}}$, and a formula $\varphi \in \mathcal{L}_{\mathcal{A}}$, a set of formulas $\Gamma \subseteq \mathcal{L}_{\mathcal{A}}$,

(1) We say φ is valid in $\mathcal{M}$ iff for every valuation function ν, we have $(\mathcal{M}, \nu) \models \varphi$. We use $\mathcal{M} \models \varphi$ to express such relation, and say $\mathcal{M}$ is a model of φ

(2) We say Γ is valid in $\mathcal{M}$, iff every formula in Γ is valid in $\mathcal{M}$. We use $\mathcal{M} \models \Gamma$ to express this relation, and say $\mathcal{M}$ is a model of Γ.

Note that if Γ is a set of sentences, then Γ is satisfiable iff it has a model.

Definition 2.18. Theory

Given a First order language $\mathcal{L}_{\mathcal{A}}$, a set of sentences $T \subseteq \mathcal{L}_{\mathcal{A}}$ is called a (First order) *theory*. If a structure $\mathcal{M}$ satisfies T, we then call $\mathcal{M}$ a model of the theory T, denote as $\mathcal{M} \models T$.

3 Completeness of First order Logic

3.1 Basic Concepts

As in the propositional logic, where we use tautologies to enrich our deduction system, here in first order logic, we try to find some thing like tautologies, called *universally valid* formulas.

Definition 3.1. A formula $\varphi \in \mathcal{L}_{\mathcal{A}}$ is said to be *universally valid* iff for every $\mathcal{L}_{\mathcal{A}}$-structure $\mathcal{M}$, we have $\mathcal{M} \models \varphi$. We denote it as $\models \varphi$.

A set of formulas $\Gamma \subseteq \mathcal{L}_{\mathcal{A}}$ is said to be universally valid iff every formula in it is valid. We denote it as $\models \Gamma$.

For example, $(x_0 \hat{=} x_0)$ is an universally valid formula.

From now on, we'll write $\mathcal{L}$ instead of $\mathcal{L}_{\mathcal{A}}$ to express a specific first order language.

We also have some ideas:

Definition 3.2. Given $\Gamma \subset \mathcal{L}$, $\varphi \in \mathcal{L}$, we say φ is a logical consequence of Γ, if for every structure $\mathcal{M}$ and every valuation function ν, if $(\mathcal{M}, \nu) \models \Gamma$, then $(\mathcal{M}, \nu) \models \varphi$.

Definition 3.3. Given $\Gamma \subset \mathcal{L}$, $\varphi \in \mathcal{L}$, we say φ is a logically independent from Γ, if $\Gamma \not\models \varphi$ and $\Gamma \not\models (\neg\varphi)$.

It seems impossible to verify whether a formula φ is logical consequence of Γ or not, since we may have to verify all possibilities on first order structures. We need an effective method to do determine.

3.2 Deduction System

In this part, we introduce a deduction system of first order logic.

Definition 3.4. Consider a set of formulas $\mathbb{L} \subseteq \mathcal{L}$, which is the smallest set satisfying all the following closure properties:

(1) Logical tautologies. Assume A, B, C are formulas, then the following are all in $\mathbb{L}$:

-$(A \to (B \to C) \to ((A \to B) \to (A \to C)))$

-$(A \to A)$

-$(A \to (B \to A))$

-$(A \to ((\neg A) \to B))$

-$(((\neg A) \to A) \to A)$

-$((\neg A) \to (A \to B))$

-$(A \to ((\neg B) \to (\neg(A \to B))))$

(2) Specialization principle. Assume φ is a formula, τ is a term, and τ can substitute x_i in φ, then the following formula is in $\mathbb{L}$:

$((\forall x_i \varphi) \to \varphi(x_i; \tau))$

(3) Distributive law for universal quantifiers. Assume A and B are formulas, then the following formula is in $\mathbb{L}$:

$((\forall x_i (A \to B)) \to ((\forall x_i A) \to (\forall x_i B)))$

(4) Introduction law of irrelevant quantifiers. Assume φ is a formula, x_i is not a free variable of φ, then the following formula is in $\mathbb{L}$:

$(\varphi \to (\forall x_i \varphi))$

(5) Law of generalization. If $\varphi \in \mathbb{L}$, then $(\forall x_i \varphi) \in \mathbb{L}$.

(6) Law of constancy. $(x_i \hat{=} x_i) \in \mathbb{L}$ for every variable symbol x_i.

(7) Equivalence law. Assume A and B are two formulas, and x_j can substitute x_i in both A and B. If we substitute x_i by x_j in A and B, we get two identical formulas, then the following formula is in $\mathbb{L}$:

$((x_i \hat{=} x_i) \to (A \to B))$

All formulas in $\mathbb{L}$ are universally valid.

With those valid materials, we introduce the concept of *proof* and *theorem*:

Definition 3.5. Given a set $\Gamma \subseteq \mathcal{L}$, we call a sequence of formulas $\langle B_1, ..., B_n \rangle$ a *proof* of Γ iff for every B_i, $1 \leq i \leq n$, one of the following holds:

(1) $B_i \in \Gamma \cup \mathbb{L}$.

(2) There exists $1 \leq j, k < i$, such that B_k is just $(B_j \to B_i)$.

Definition 3.6. Given a set $\Gamma \subseteq \mathcal{L}$ and a formula φ, we say φ is a *theorem* of Γ, or $\Gamma \vdash \varphi$, iff there exists a proof $\langle B_1, ..., B_n \rangle$ of Γ, with $\varphi = B_n$.

As in the propositional logic, we have concept of consistency:

Definition 3.7. A set of formulas Γ is said to be consistent, iff for arbitrary formula φ, if $\Gamma \vdash \varphi$, then $\Gamma \nvdash (\neg\varphi)$

A set of formulas Γ is said to be inconsistent iff it's not consistent.

3.3 Four Effective Theorems

In this part, we introduce four effective theorems: Deduction Theorem, Generalization Theorem, Theorem on Constants, Equality Theorem. Those theorems are useful in the proof of completeness theorem.

Theorem 3.1. Deduction Theorem

$\Gamma \cup \{\varphi\} \vdash \psi$ iff $\Gamma \vdash (\varphi \to \psi)$

And it's corollary:

Corollary 3.1. If $\Gamma \cup \{(\neg\varphi)\} \vdash \varphi$, then $\Gamma \vdash \varphi$.

The two results above are actually almost the same as that in propositional logic.

Moreover, we have:

Theorem 3.2. Generalization Theorem

If $\Gamma \vdash \varphi$, and if x_i is not a free variable of any formula in Γ, then $\Gamma \vdash (\forall x_i \varphi)$.

We need a generalized concept of substitution:

Definition 3.8. Given a formula φ, a constant element c_i and a variable x_j, we say x_j can substitute c_i in φ, iff any occurrence of c_i in φ is not in range of any occurrence of $\forall x_j$ in φ.

Theorem 3.3. Theorem on Constants

Assume (1) $\Gamma \vdash \varphi$, (2) c_i occurs nowhere in any Γ and (3) x_j has no free occurrence in φ and x_j can substitute c_i in φ.

Then $\Gamma \vdash (\forall x_j \varphi(c_i; x_j))$, and $(\forall x_j \varphi(c_i; x_j))$ is proved by a sequence $\langle B_1, ..., B_n \rangle$ satisfying

1. For $1 \leq m \leq n$, c_i doesn't occur in B_m.
2. For $1 \leq m < n$, if c_k appears in B_m, then

(a) either c_k occurs in $(\forall x_j \varphi(c_i; x_j))$,

(b) or c_k occurs in some formula of Γ.

To prove it, we still need 2 more lemmas:

Lemma 3.1. Change Quantifiers

Given $\varphi \in \mathcal{L}$, x_i can substitute x_j in φ, and x_i has no free occurrence in $(\forall x_j \varphi)$, then $\vdash ((\forall x_j \varphi) \to (\forall x_i \varphi(x_j; x_i)))$.

Lemma 3.2. Lemma on Constants

Given $\Gamma \subseteq \mathcal{L}$, and c_i occurs nowhere in Γ. If $\langle B_1, ..., B_n \rangle$ is a Γ-proof, and x_j appears nowhere in Γ. Then $\langle B_1(c_i; x_j), ..., B_n(c_i; x_j) \rangle$ is still a Γ-proof.

Finally, the last theorem:

Theorem 3.4. Equality Theorem

Assume $\varphi \in \mathcal{L}$ has no quantifiers, $\sigma_1, ..., \sigma_n$ and $\tau_1, ..., \tau_n$ are two collection of terms. Given variables $x_{m_1}, ..., x_{m_n}$, if each of them occurs nowhere in σ_i, τ_i for every $1 \leq i \leq n$, then we have

$\{\tau_i \hat{=} \sigma_i | 1 \leq i \leq n\} \cup \{\varphi(x_{m_1}, ..., x_{m_n}; \tau_1, ..., \tau_n)\} \vdash \varphi(x_{m_1}, ..., x_{m_n}; \sigma_1, ..., \sigma_n)$

3.4 Completeness

In this part we discuss the most important part: completeness of first order logic.

Theorem 3.5. Soundness

If $\Gamma \vdash \varphi$, then $\Gamma \models \varphi$.

Theorem 3.6. If Γ is satisfiable, then Γ is consistent.

Proof. Otherwise assume Γ is inconsistent, then there is a formula φ such that $\Gamma \vdash \varphi$ and $\Gamma \vdash (\neg\varphi)$. By Soundness Theorem, we have $\Gamma \models \varphi$ and $\Gamma \models (\neg\varphi)$.

On the other hand, there is a structure $\mathcal{M}$ and a valuation ν such that $(\mathcal{M}, \nu) \models \Gamma$, then $(\mathcal{M}, \nu) \models \varphi$ and $(\mathcal{M}, \nu) \models\models (\neg\varphi)$, contradicts. So Γ must be consistent. $\square$

To prove the other hand, we need much more work. We start from some similar concepts as in propositional logic.

Definition 3.9. Maximally Consistency

A consistent set $\Gamma \subseteq \mathcal{L}$ is said to be *maximally consistent* iff for every formula $\varphi \in \mathcal{L}$, if $\Gamma \cup \{\varphi\}$ is consistent, then $\varphi \in \Gamma$.

For example, given a structure $\mathcal{M} = (M, I)$ and a valuation ν, the set $\{\varphi \in \mathcal{L} | (\mathcal{M}, \nu) \models \varphi\}$ is maximally consistent.

Maximally consistent sets have good properties:

Lemma 3.3. Given a maximally consistent set $\Gamma \subseteq \mathcal{L}$, we have

1. If $\Gamma \vdash \varphi$, then $\varphi \in \Gamma$.
2. Either $\varphi \in \Gamma$ or $(\neg\varphi) \in \Gamma$.
3. If $\varphi \in \Gamma$, and $(\varphi \to \psi)$, then $\psi \in \Gamma$.

Different from the case in propositional logic, we still need another important property: *Henkin* properties.

Definition 3.10. A set $\Gamma \subseteq \mathcal{L}$ is said to have *Henkin* properties, iff for every formula φ and variable x_i, if $(\exists x_i \varphi) \in \Gamma$, then there must be a constant element symbol c_j, such that $\varphi(x_i; c_j) \in \Gamma$.

Now we have a powerful Theorem:

Theorem 3.7. Satisfiability Theorem

A maximally consistent set $\Gamma \subseteq \mathcal{L}$ with Henkin property is satisfiable.

With such theorem, to answer the final question on the satisfiability of arbitrary consistent set, we need other results: We need to extend an arbitrary consistent set into a maximally consistent set with Henkin property. We proceed as the following.

We need enough constant element symbol. Actually we can save countably infinite many constant element symbols with a little trick:

Definition 3.11. Formula Transformation

Given arbitrary formula φ, we'd better think φ is a map from some positive integer $(n+1)$ to the set of of all logical symbols. Then we define another formula φ^*, we may call this formula the adjoint formula of φ. For every m, $0 \leq m \leq n$, we define

$$\varphi^*(m) = \begin{cases} c_{2i} & \text{if } \varphi(m) \text{ is a constant element symbol, and } c_i = \varphi(m) \\ \varphi(m) & \text{if } \varphi(m) \text{ is not a constant element symbol} \end{cases} \tag{2}$$

It's not hard to see that φ^* is also a formula, and it has no constant element symbol with odd index.

Definition 3.12. Given $\Gamma \subseteq \mathcal{L}$, define:

$\Gamma^* \{\varphi^* | \varphi \in \Gamma\}$.

Note that Γ and Γ^* have strong relations:

Lemma 3.4. If Γ is consistent, then Γ^* is consistent also.

Proof. For every $\psi^* \in \Gamma^*$, we can get a formula $\psi \in \Gamma$, simply substitute constant element symbols with index $2n$ in ψ by constant element symbols with index n.

If Γ^* is inconsistent, then $\Gamma^* \vdash (\neg(x_0 \hat{=} x_0))$, so there is a Γ^*-proof of it: $\langle B_1^*, ..., B_n^*$. Then we know that there are unique formulas in Γ, which are adjoints of those B_i^*'s, we denote them as B_i.

It's not hard to show that $\langle B_1, ..., B_N \rangle$ is Γ-proof of $(\neg(x_0 \hat{=} x_0))$. Which means $\Gamma \vdash (\neg(x_0 \hat{=} x_0))$, this contradicts the assumption that Γ is consistent.

Thus Γ^* must be consistent. □

Lemma 3.5. If Γ^* is satisfiable, then Γ is satisfiable.

Proof. Assume structure $\mathcal{M} = (M, I*)$ together with valuation function ν, satisfies Γ^*.

Then we know the set of all non-logical symbols used in Γ, denote as $\mathcal{A}$, and Γ^*'s as $\mathcal{A}^*$.

We know that $\mathcal{A} = \{F_n, P_n | F_n, P_n \in A^*\} \cup \{c_k | c_{2k} \in A^*\}$.

So we on the same domain M, we define another interpretation I:

$I(c_k) = I^*(c_{2k})$,

$I(F_n) = I^*(F_n)$,

$I(P_n) = I^*(P_n)$.

Then it's not hard to show that(by induction), for arbitrary valuation function mu:

$\bar(\nu)(\tau) = \bar(\nu)^*(\tau^*)$(we use notations sloppily here, but readers should know what do those means).

So by induction again, we have $((M, I), \mu) \models \Gamma$ iff $((M, I^*), \mu) \models \Gamma^*$.

Since $((M, I^*), \nu) \models \Gamma^*$, we have $((M, I), \nu) \models \Gamma$. Therefore Γ is satisfiable. □

So we may consider Γ^* instead of Γ, to save more constant element symbols.

Theorem 3.8. Extension Theorem

Assume Γ is consistent, and there are infinitely many constant element symbols that occurs nowhere in Γ, then Γ must a subset of some maximally consistent set with Henkin property.

With results above, we can prove the completeness theorem:

Theorem 3.9. Completeness Theorem

If Γ is consistent, then it is satisfiable.

Proof. We consider Γ^*, we know constant elements with odd indices will not appear in it, so by Extension Theorem, there must be a maximally consistent set Γ_1 with Henkin property, such that $\Gamma^* \subseteq \Gamma_1$.

By Satisfiability Theorem, we know that Γ_1 is satisfiable, then Γ^* is satisfiable also. By lemma above, we know Γ is satisfiable, completing the proof. □

Corollary 3.2. Completeness

Assume $\Gamma \vdash \varphi$, then $\Gamma \vdash \varphi$.

Proof. Assume $\Gamma \nvdash \varphi$, then $\Gamma \cup \{(\neg\varphi)\}$ is consistent, so by completeness theorem, we know $\Gamma \cup \{(\neg\varphi)\}$ is consistent. Then there is a structure $\mathcal{M}$ and a valuation function ν, such that $(\mathcal{M}, \nu) \models \Gamma \cup \{(\neg\varphi)\}$.

So $(\mathcal{M}, \nu) \models (\neg\varphi)$, $(\mathcal{M}, \nu) \models \Gamma$, and since $\Gamma \models \varphi$, we have $(\mathcal{M}, \nu) \models \varphi$, which is impossible.

Therefore $\Gamma \vdash \varphi$. □

As in the propositional logic, we have compactness:

Theorem 3.10. Γ is consistent iff every finite subset of it is consistent.

Proof. Use finiteness of proofs. □

4 Discussions

Based on completeness of first order logic, we discuss some specific theories, models and so on.

4.1 Non-standard Model of Arithmetic

4.2 Completeness of Theories