Mathematical Foundations of computer science

Lecture 8: Proofs in FO logic, soundness and completeness

Speaker: Toru Takisaka

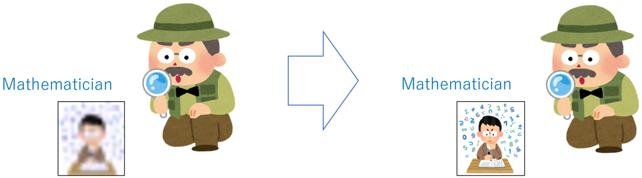
April 23, 2024

Recap

Current goal: formalize Mathematician with higher resolution

- We construct atomic propositions from terms / predicates
- We defined structures that specify the meaning of formulas
- We defined the truth of sentences under a given structure
  - Sentence = formula with no free variables

- Today: proofs in FO logic, soundness/completeness



A minor correction on formulas' definition

We defined atomic predicate formulas as follows;

- An expression $P(t_1, \ldots, t_k)$, where t_1, …, t_k are terms and $P \in L$ is a k-ary predicate symbol;
- An expression $(t_1 = t_2)$, where t_1 and t_2 are terms.

Let's say $\top$, $\bot$ are also atomic formulas (as we do in propositional logic).

However, both of old and new definitions work in a same way; in the old one, we can pick any tautology and contradiction, and call them $\top$ and $\bot$, respectively.

Recap: truth of sentences

We can determine the truth of a formula (given a structure) when it has no free variable. Such a formula is called a **sentence**.

- Consider a formula $\mathit{IsOdd}(n)$ with the structure as given in p10. Thus it could read "a natural number *n* is odd".

  $\rightarrow$ We cannot determine the truth of this formula, as it depends on what *n* is.

- Consider a formula $\exists n.\mathit{IsOdd}(n)$ with the same structure. Thus it could read "there exists an odd natural number *n*".

  $\rightarrow$ the truth of this formula should be determined, because the meaning of *n* is *bound* by the prefix "there exists".

n in the first formula is **free**; *n* in the second is **bound**.

Recap: truth of sentences

Suppose we would like to define the truth of sentences written in a language L under a structure $\mathcal{A}$, whose domain is A.

To this end, we augment L by adding a new constant $\overline{a}$ for each $a \in A$. By this, we can instantiate a formula $\varphi(x)$ into $\varphi(\overline{a})$, which claims a property about a particular object a.

- You can write formulas that claim e.g. "1 is odd" "2 is odd", not only "n is odd" (where n is a variable)

Let $L_{\mathcal{A}}$ be such an extension. We also extend $\mathcal{A}$ so that it assigns a to $\overline{a}$, for each $a \in A$. Let $\mathcal{A}^*$ be such an extension.

Truth of sentences and formulas, formally

Let $SENT^L$ be the set of all sentences written in L.

Recall a structure $\mathcal{A}$ specifies the interpretation of any ground term t, denoted by $t^{\mathcal{A}}$. Also, $P^{\mathcal{A}}$ is the interpretation of P by $\mathcal{A}$.

Definition

The set $Th(\mathcal{A})$ of sentences written in $L_{\mathcal{A}}$ that are true under $\mathcal{A}^*$ is the smallest subset of $SENT^{L_{\mathcal{A}}}$ that satisfy:

- $\top \in Th(\mathcal{A})$ and $\bot \notin Th(\mathcal{A})$;
- if $\varphi \equiv P(t_1, \ldots, t_k)$, then $\varphi \in Th(\mathcal{A})$ iff $(t_1^{\mathcal{A}^*}, \ldots, t_k^{\mathcal{A}^*}) \in P^{\mathcal{A}^*}$;
- if $\varphi \equiv t_1 = t_2$, then $\varphi \in Th(\mathcal{A})$ iff $t_1^{\mathcal{A}^*} = t_2^{\mathcal{A}^*}$;
- if $\varphi \equiv \neg\psi$, then $\varphi \in Th(\mathcal{A})$ iff $\psi \notin Th(\mathcal{A})$;
- if $\varphi \equiv (\psi_1 \vee \psi_2)$, then $\varphi \in Th(\mathcal{A})$ iff $\psi_1 \in Th(\mathcal{A})$ or $\psi_2 \in Th(\mathcal{A})$;
- if $\varphi \equiv (\psi_1 \wedge \psi_2)$, then $\varphi \in Th(\mathcal{A})$ iff $\psi_1 \in Th(\mathcal{A})$ and $\psi_2 \in Th(\mathcal{A})$;
- if $\varphi \equiv \exists x.\varphi(x)$, then $\varphi \in Th(\mathcal{A})$ iff $\varphi(\overline{a}) \in Th(\mathcal{A})$ for some $a \in A$;
- if $\varphi \equiv \forall x.\varphi(x)$, then $\varphi \in Th(\mathcal{A})$ iff $\varphi(\overline{a}) \in Th(\mathcal{A})$ for any $a \in A$.

Truth of sentences and formulas, formally

Let φ be a formula whose free variables are $x_1, \dots, x_k$. We call the sentence $\forall x_1 \dots \forall x_k.\varphi$ the **universal quantification** of φ.

Definition (truth of first-order sentences)

We say a sentence φ written in L is true under $\mathcal{A}$, written as $\mathcal{A} \models \varphi$, when $\varphi \in \mathit{Th}(\mathcal{A})$.
For a general formula φ, we say φ is true under $\mathcal{A}$ if its universal quantification is in $\mathit{Th}(\mathcal{A})$.

Notions related to the truth of formulas

We say φ is a **tautology** if $\mathcal{A} \models \varphi$ for any structure $\mathcal{A}$.
We say φ is a **contradiction** if $\mathcal{A} \not\models \varphi$ for any structure $\mathcal{A}$.

We say a set Σ of formulas **logically implies** a formula φ, written as $\Sigma \models \varphi$, if any structure $\mathcal{A}$ that makes $\mathcal{A} \models \psi$ for every $\psi \in \Sigma$ makes $\mathcal{A} \models \varphi$ also true.

We call a structure $\mathcal{A}$ a **model** of a set Σ of formulas if $\mathcal{A} \models \psi$ holds for every $\psi \in \Sigma$.

On logical implication in predicate logic

In propositional logic, if Σ logically implies $\varphi \to \psi$, then $\Sigma \cup \{\varphi\}$ logically implies ψ, and vice versa. In predicate logic, the latter holds when φ is a sentence.

Lemma

Let Σ be a set of formula, and φ, ψ are formulas. Then we have:

1. *If $\Sigma \cup \{\varphi\} \models \psi$ and φ is a sentence, then $\Sigma \cup \{\varphi\} \models \psi$.*
2. *If $\Sigma \cup \{\varphi\} \models \psi$, then $\Sigma \cup \{\varphi\} \models \psi$.*

Non-example: $\varphi(x) \models \forall y.\varphi(y)$ generally does not imply $\varphi(x) \to \forall y.\varphi(y)$ is a tautology.

- The latter statement means $\forall x.(\varphi(x) \to \forall y.\varphi(y))$ is a tautology, which can be false (for example, imagine the statement "if 1 is odd, then any natural number is odd").

Examples of models

Consider a structure $\mathcal{A} = (H; \mathit{IsProf}_{\mathcal{A}}, \mathit{IsInChina}_{\mathcal{A}}, \mathit{Over}20_{\mathcal{A}})$ for a language $L = (\mathit{IsProf}, \mathit{IsInChina}, \mathit{Over}20)$ such that

- H is the set of humans (say, who were alive at a certain timestamp),
- $\mathit{IsProf}_{\mathcal{A}} \subseteq H$ is the set of professors in H (say, at that timestamp),
- $\mathit{IsInChina}_{\mathcal{A}} \subseteq H$ is the set of residents in China in H, and
- $\mathit{Over}20_{\mathcal{A}} \subseteq H$ is the set of humans in H that are over 20 years old.

Suppose any professor in H who lives in China is over 20years old. Then $\mathcal{A}$ models $\Sigma = \{\varphi\}$, where

$$\varphi \equiv \forall x. \Big((\mathit{IsProf}(x) \wedge \mathit{IsInChina}(x)) \rightarrow \mathit{Over}20(x) \Big).$$

Examples of models

Consider a structure $\mathcal{A} = (\mathbb{N}; 0, S)$, where

- $\mathbb{N}$ is the set of natural numbers, and
- $S : \mathbb{N} \to \mathbb{N}$ is the successor function (i.e., $S(n) = n + 1$).

Also let Σ consists of the following sentences*, which are called **axioms of natural numbers**:

1. $\forall x.\forall y.(S(x) = S(y) \to x = y)$
2. $\forall x.\neg(S(x) = 0)$
3. $\forall x.\Big(\neg(x = 0) \to \exists y.(x = S(y))\Big)$

Then $\mathcal{A}$ is a model of Σ.

*) precisely speaking, we have a language $L = (\overline{0}, \overline{S})$ in mind and we should write these axioms by $\overline{0}$, $\overline{S}$ instead of 0, S. But the meaning should be clear.

Now we formalize the notion of proofs in predicate logic. Recall, in propositional logic, we needed the following ingredients:

- Axioms, which are formulas that we admit to be "generally true";
- Inference rules, which allows us to infer a new formula out of certain formulas.

Then we defined a proof as a sequence of formulas from axioms and assumptions to the target formula, connected by inference rules.

We define proofs in predicate logic in a similar way.

Propositional Axioms

As axioms, We first register similar formulas as axioms in propositional logic. Namely, for any formula φ, ψ and γ, the following formulas are axioms.

1. $\top$.
2. $\varphi \to (\varphi \vee \psi)$.
3. $\varphi \to (\psi \vee \varphi)$.
4. $(\varphi \,\&\, \psi) \to \varphi$.
5. $(\varphi \,\&\, \psi) \to \psi$.
6. $(\varphi \to (\psi \to (\varphi \,\&\, \psi)))$.
7. $(\neg\varphi \to (\neg\psi \to \neg(\varphi \vee \psi)))$.
8. $(\varphi \to (\psi \to \gamma)) \to ((\varphi \to \psi) \to (\varphi \to \gamma))$.
9. $(\varphi \to (\neg\varphi \to \bot))$.
10. $(\neg\varphi \to \bot) \to \varphi)$.

Equality Axioms

We have the equality symbol "$=$", which should have a fixed interpretation. To this end, we register the following formulas as axioms.

1. $x = x$.
2. $x = y \to y = x$.
3. $(x = y \;\&\; y = z) \to (x = z)$.
4. $(x_1 = y_1 \;\&\; \ldots \;\&\; x_k = y_k) \to (P(x_1, \ldots, x_k) \leftrightarrow P(y_1, \ldots, y_k))$, where P is a predicate symbol of arity k.
5. $(x_1 = y_1 \;\&\; \ldots \;\&\; x_k = y_k) \to f(x_1, \ldots, x_k) = f(y_1, \ldots, y_k)$, where f is a operation symbol of arity k.

Quantifier axioms

We have new logical connectives that do not appear in propositional logic, namely $\exists x$ and $\forall x$. We also register the following formulas as axioms. Here, t is any term.

1. $\varphi(t/y) \to \exists y \varphi$.
2. $\forall y \varphi(y) \to \varphi(t/y)$.
3. $\forall x \varphi(x) \to \neg\exists x \neg\varphi(x)$.
4. $\neg\exists x \neg\varphi(x) \to \forall x \varphi(x)$.
5. $\exists x \varphi(x) \to \neg\forall x \neg\varphi(x)$.
6. $\neg\forall x \neg\varphi(x) \to \exists x \varphi(x)$.

Axioms are sound

Lemma (soundness of axioms)

All axioms are tautologies, that is, any axiom φ and structure $\mathcal{A}$ satisfy $\mathcal{A} \models \varphi$.

Proof.

One proves this by going through the axioms. As an example, consider the axiom scheme $\gamma \equiv \varphi \to (\varphi \vee \psi)$. Let $x_1, \ldots, x_m$ be all the variables in φ and ψ. We let $x_1 = \overline{a_1}, \ldots, x_m = \overline{a_m}$ and evaluate $\mathcal{A} \models \varphi(\overline{a_1}, \ldots, \overline{a_m})$ and $\mathcal{A} \models \psi(\overline{a_1}, \ldots, \overline{a_m})$. Substitute these truth values into $\varphi \to (\varphi \vee \psi)$, and compute the value $\mathcal{A} \models \gamma(\overline{a_1}, \ldots, \overline{a_m})$. Since $\varphi \to (\varphi \vee \psi)$ is a propositional axiom, we know that it is always evaluated to **true** independent of values of $\mathcal{A} \models \varphi(\overline{a_1}, \ldots, \overline{a_m})$ and $\mathcal{A} \models \psi(\overline{a_1}, \ldots, \overline{a_m})$. Hence, the value of $\mathcal{A} \models \gamma(\overline{a_1}, \ldots, \overline{a_m})$ is **true** for any $\mathcal{A}$ and $a_1, \ldots, a_m$; therefore, we have $\mathcal{A} \models \forall x_1 \ldots \forall x_m . \gamma(x_1, \ldots, x_m)$. $\square$

Inference rules

The first order logic uses the following inference rules:

1. The *Modus Ponens (MP) rule*: From φ and $\varphi \to \psi$ infer ψ.
2. The *Generalisation (G)* rules. Let x be not free in φ.
    1. From $\varphi \to \psi$ infer $\varphi \to \forall x \psi$.
    2. From $\psi \to \varphi$ infer $\exists x \psi \to \varphi$.

Lemma (soundness of inference rules)

Let $\mathcal{A}$ be an algebraic structure.

1. *If $\mathcal{A} \models \varphi$ and $\mathcal{A} \models \varphi \to \psi$, then $\mathcal{A} \models \psi$.*
2. *Let x be a variable not free in φ.*
    1. *If $\mathcal{A} \models \varphi \to \psi$ then $\mathcal{A} \models \varphi \to \forall x \psi$.*
    2. *If $\mathcal{A} \models \psi \to \varphi$ then $\mathcal{A} \models \exists x \psi \to \varphi$.*

Item 2 in the lemma does not hold if we allow x to be free in φ; see the non-example in p9.

Definition of proof in the FO logic

From now on we fix Σ a set of sentences. Let ψ be a formula.

Definition

A **proof** of ψ from Σ is a sequence $\psi_1, \psi_2, \ldots, \psi_n$ of formulas such that $\psi_n = \psi$, and for all $k = 1, \ldots, n$, either

- ψ_k is an axiom, or
- $\psi_k \in \Sigma$, or
- there are $i, j < k$ such that φ_k is inferred from ψ_i and ψ_j via Modus Ponens rule, or ψ_k is inferred from φ_i via the generalisation rule.

Call n the length of the proof. If there is a proof of ψ from Σ, then write this $\Sigma \vdash \psi$. If $\Sigma = \emptyset$, then write this $\vdash \psi$.

Soundness theorem

Theorem (Soundness Theorem)

If $\Sigma \vdash \varphi$ *then* $\Sigma \models \varphi$*.*

Proof.

Let $\mathcal{A}$ be a model of Σ, and let $\varphi_1, \ldots, \varphi_k$ be a proof of φ from Σ. Then by soundness of axioms and inference rules, one can inductively show $\mathcal{A} \models \varphi_j$ for each $j \in \{1, \ldots, k\}$ (observe each $\psi \in \Sigma$ satisfies $\mathcal{A} \models \psi$). As $\varphi_k \equiv \varphi$ by definition of a proof, we have $\mathcal{A} \models \varphi$; hence we have $\Sigma \models \varphi$. $\square$

Completeness theorems

Definition

We say a set Σ of formulas is **inconsistent** if $\Sigma \vdash \bot$. Otherwise, we say that Σ is consistent.

Theorem (The second completeness theorem)

A set Σ of sentences is consistent if and only if Σ has a model.

Theorem (The first completeness theorem)

Let Σ be a set of sentences and φ be a formula. Then $\Sigma \vdash \varphi$ if and only if $\Sigma \models \varphi$.