Mathematical Foundations of computer science

Lecture 9: Introduction to Part 3 (incompleteness theorems)

Speaker: Toru Takisaka

April 25, 2024

Recap: what we have done so far

2nd goal: formalize a “mathematician” with higher resolution

- Analysis target of Mathematician $\approx$ Formulas
  - Math assertions $\to$ Formulas, made of terms/predicates
  - Truth of assertions are determined by fixing a structure
- Mathematician $\approx$ those who write proofs of formulas
  - Tools: axioms, assumptions, inference rules
  - Proof $\to$ seq. of formulas from axioms/assump. to target, connected by inference rules
- Properties of Mathematician
  - Soundness theorem: proved formulas are “true”
  - Completeness theorem: “true” formulas can be proved

Recap: what we have done so far

There were two important notions, $\Sigma \models \varphi$ and $\Sigma \vdash \varphi$.

- $\Sigma \models \varphi$ (Σ logically implies φ) talks about the nature of the truth of formulas, which Mathematician wants to analyze. It is a semantic notion (it is about "the meaning" of formulas).
- $\Sigma \vdash \varphi$ (φ has a proof from Σ) talks about the possibility of inferring φ from Σ, which Mathematician attempts. Recall inference is a syntactic operation (i.e., inference rules can be described at the grammatical level).

Excercise: Prove $\varphi \models \varphi \vee \psi$ and $\varphi \vdash \varphi \vee \psi$, and observe how different their proofs are.

Recap: what we have done so far

Soundness/completeness theorems say these notions coincide

(in predicate logic, we typically assume Σ is a set of sentences, while φ is any formula).

Theorem (Soundness of proof structure)

$\Sigma \vdash \varphi$ *implies* $\Sigma \models \varphi$.

The theorem roughly claims "correctness" of Mathematician: a proved formula under assumptions Σ is true whenever Σ is.

Theorem ((First) completeness of proof structure)

$\Sigma \models \varphi$ *implies* $\Sigma \vdash \varphi$.

The theorem roughly claims the "capability" of Mathematician: They can prove *any* φ from Σ whenever Σ logically implies φ.

Seeing completeness theorem, you may feel:

- We are all done now, don't we?
  We know our "mathematician" can prove any "true" formula, doesn't this mean he is capable enough?

Now let's try to understand what it means exactly...

One headsup here (not serious):

Try to count how many times we use soundness/completeness theorems in the rest of the lecture today. You'll realize how fundamental they are. :)

What is our interest exactly?

Rather often, we would like to know if a formula φ is true or not under a particular structure $\mathcal{A}$.

- In previous lectures, we considered a sentence

$$\forall x. \Big((\mathit{IsProf}(x) \wedge \mathit{IsInChina}(x)) \rightarrow \mathit{Over20}(x) \Big).$$

  This means "Any professor in China is over 20years old" under a particular structure (e.g., the canonical one we constructed before), but not under an arbitrary structure.

What is our interest exactly?

Rather often, we would like to know if a formula φ is true or not under a particular structure $\mathcal{A}$.

- In previous lectures, we also considered a sentence

$$\exists x.\forall y.\Big(prime(y) \land prime(y+2) \to y \leq x\Big).$$

This means the *twin prime conjecture* under the standard model of arithmetic, but not under an arbitrary structure.

What is our interest exactly?

Based on these observations, we could formalize our interest on the capability of “mathematician” as follows:

Fix a structure $\mathcal{A}$, and a set Σ of sentences modeled by $\mathcal{A}$.

1. For which φ such that $\mathcal{A} \models \varphi$ do we have $\Sigma \vdash \varphi$?
   (It reads: Which φ that are true under $\mathcal{A}$ can be proved from Σ?)
2. For which φ such that $\mathcal{A} \not\models \varphi$ do we have $\Sigma \vdash \neg\varphi$?
   (It reads: Which φ that are false under $\mathcal{A}$ can be disproved from Σ?)

What is our interest exactly?

The best outcome is that we have either $\Sigma \vdash \varphi$ or $\Sigma \vdash \neg\varphi$ for *any* φ, i.e., we can either prove or disprove any formula from Σ.

- We require $\mathcal{A}$ models Σ (thus in particular, Σ is consistent); this is necessary for having $\Sigma \vdash \varphi \implies \mathcal{A} \models \varphi$.

Definition

We say a set Σ of sentences is **complete**[1] if either $\Sigma \vdash \varphi$ or $\Sigma \vdash \neg\varphi$ holds for any formula φ; if not, Σ is **incomplete**.

Thus we can shorten our interest as follows:

Let $\mathcal{A}$ be given; is there a complete Σ modeled by $\mathcal{A}$?

[1] in previous lectures we used this word in a bit different way; see the followup material.

A boring answer

Actually, the answer is trivially yes for any $\mathcal{A}$;
let Σ be the set of all sentences that are true under $\mathcal{A}$.

However, such a Σ is “unrealistic” as the initial knowledge of a mathematician; knowing such a Σ means that he already knows the truth of every sentence under $\mathcal{A}$.

Now which Σ are “realistic”?

Realistic = Computable

Here computer science meets logic: let's say Σ is "realistic" when a computer can "generate" Σ in a certain sense.

Definition (recursive sets of sentences; a bit informal for now)

We say a set Σ of sentences is **recursive** if there is a program that receives a sentence φ as an input, and returns

- **true** if $\varphi \in \Sigma$, and
- **false** if $\varphi \notin \Sigma$.

So let's modify our interest as follows:

Let $\mathcal{A}$ be given; is there a recursive complete Σ modeled by $\mathcal{A}$?

Let's consider the case with the weakest assumption, $\Sigma = \emptyset$.
In this case, soundness/completeness says the following:

$$\vdash \varphi \quad \text{iff} \quad \models \varphi.$$

That is, φ can be proved iff φ is true under *any* structure $\mathcal{A}$.
This means we *cannot* prove any φ that is false under some $\mathcal{A}$.

Ex.) Let $L = (\overline{0}, S, +, \times)$ (the language of arithmetic).
We cannot prove any of the following sentences from $\Sigma = \emptyset$:

$$\exists x.(x \neq \overline{0}), \quad \forall x.(x + \overline{0} = x), \quad \forall x.\forall y.(x + y = y + x).$$

How things look like when Σ is empty

When $\Sigma = \emptyset$, our "mathematician" has no ability to prove any φ such that $\mathcal{A} \models \varphi$, but $\mathcal{B} \not\models \varphi$ for some $\mathcal{B}$.

One can imagine Σ is "the initial knowledge of mathematician" (together with axioms); thus, $\Sigma = \emptyset$ means he does not know anything about $\mathcal{A}$.

So let's augment Σ by adding sentences that explain $\mathcal{A}$.

Case 1: propositional logic

Let's consider the case of propositional logic with $P = \{p, q\}$.
Recall, in propositional logic, we have a *truth assignment* $A : P \to \{\mathbf{true}, \mathbf{false}\}$ instead of a structure $\mathcal{A}$.

In this setting, there are 4 possible truth assignments $A_1 \sim A_4$:

1. $A_1(p) = \mathbf{true}$, $A_1(q) = \mathbf{true}$,
2. $A_2(p) = \mathbf{true}$, $A_2(q) = \mathbf{false}$,
3. $A_3(p) = \mathbf{false}$, $A_3(q) = \mathbf{true}$, and
4. $A_4(p) = \mathbf{false}$, $A_4(q) = \mathbf{false}$.

Now take any truth assignment, say A_2 for example.
Which propositional formula φ such that $A_2 \models \varphi$ can we prove?

Case 1: propositional logic

If $\Sigma = \emptyset$, there are many φ that cannot be proved, e.g.,

$$p, \quad \neg q, \quad p \vee q, \quad p \wedge \neg q, \quad \neg p \to \bot, \quad \ldots$$

because none of them are tautologies.

Now add formulas p and $\neg q$ to Σ; these formulas "explain" which truth assignment we are talking about (that is, A_2).

- Observe A_2 models Σ after this addition. If not, then it means Σ "wrongly explains" A_2 (e.g., $\neg q$ "wrongly explains" A_2).

Now we see, from $\Sigma = \{p, \neg q\}$, we can prove all formulas above (because they are all logically implied from Σ).

Case 1: propositional logic

In fact, A_2 is the unique model of $\Sigma = \{p, \neg q\}$ (A_1, A_3, A_4 do not model Σ). In this sense, Σ *determines* the truth assignment we consider, that is, A_2.

In such a case, completeness theorem tells us we can either prove or disprove *any* φ from Σ.

Corollary

Suppose Σ has a unique model A. Then for any φ, we have

1. *If $A \models \varphi$ then $\Sigma \vdash \varphi$,*
2. *If $A \not\models \varphi$ then $\Sigma \vdash \neg\varphi$.*

Proof.

By uniqueness we have $A \models \varphi$ iff $\Sigma \models \varphi$; thus by completeness theorem, $\Sigma \vdash \varphi$. If $A \not\models \varphi$ then $A \models \neg\varphi$; thus $\Sigma \vdash \neg\varphi$ by (1). $\square$

Case 1: propositional logic

We can do the similar argument for any set P of atomic proposition, and its truth assignment A.

Also, if P is finite, then Σ that we construct is recursive.

- informally speaking, consider a program that reads Σ, and returns **true** if it finds the input formula in Σ, or returns **false** otherwise.

Therefore, for finite P, we always have the desired Σ for any A (i.e., a recursive complete Σ that models A).

Case 2: structures with finite domains

Now let's consider the case of the "teachers' age problem".
There, we had the following structure $\mathcal{A}$ in mind:

$\mathcal{A} = (H; \mathit{IsProf}_{\mathcal{A}}, \mathit{IsInChina}_{\mathcal{A}}, \mathit{Over}20_{\mathcal{A}})$, where

- 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.

By which sentences can we "explain" this structure $\mathcal{A}$?
More desirably, can we "determine" $\mathcal{A}$ by certain sentences?

Case 2: structures with finite domains

Let's begin with "explaining" the domain $H = \{h_1, \ldots, h_n\}$.
For this, we use the augmented language $L_{\mathcal{A}}$, that is,

$$L_{\mathcal{A}} = (\mathit{IsProf}, \mathit{IsInChina}, \mathit{Over}20, \overline{h_1}, \ldots, \overline{h_n}).$$

Let's add the following sentences to Σ:

- $\forall h.(h = \overline{h_1} \vee \ldots \vee h = \overline{h_n})$,
- $\neg(\overline{h_m} = \overline{h_{m'}})$ for each $m \neq m'$.

After this, a structure $\mathcal{A}$ is a model of Σ if and only if its domain has exactly n elements in it (check it!).
In this sense, now Σ "determines" the domain of its model.

Case 2: structures with finite domains

Now let's also "explain" each predicate symbol in L.
For each $k \in \{1, \ldots, n\}$, add the following sentence to Σ:

- $\mathit{IsProf}(\overline{h_k})$ if $h_k \in \mathit{IsProf}_{\mathcal{A}}$,
- $\neg\mathit{IsProf}(\overline{h_k})$ if $h_k \notin \mathit{IsProf}_{\mathcal{A}}$.

Also add sentences for *IsInChina* and *Over*20 in a similar way.

Now $\mathcal{A}$ models Σ (precisely speaking, $\mathcal{A}^*$ in Lecture 7 models Σ);
furthermore, Σ "determines" $\mathcal{A}$ *up to isomorphism*.

- That is, any model $\mathcal{B}$ of Σ is **isomorphic** to $\mathcal{A}$, which roughly means $\mathcal{A}$ and $\mathcal{B}$ only differs in the label of their domains (See Definition 13 in the textbook).

Case 2: structures with finite domains

When Σ determines $\mathcal{A}$ up to isomorphism, we have a similar result as the first example (proof is similar to Case 1).

Corollary

Suppose Σ has a model $\mathcal{A}$, and any model of Σ is isomorphic to $\mathcal{A}$. Then for any φ, we have

1. *If $\mathcal{A} \models \varphi$ then $\Sigma \vdash \varphi$,*
2. *If $\mathcal{A} \not\models \varphi$ then $\Sigma \vdash \neg\varphi$.*

More generally, for any $\mathcal{A}$ with a finite domain, we can find Σ as above, which is complete and modeled by $\mathcal{A}$.
Also, this Σ is finite, so it is recursive.

Therefore, for $\mathcal{A}$ with a finite domain, we have a desired Σ.

Structures with infinite domains?

Now how about structures with infinite domains? For example, the **standard model of arithmetic** $\mathcal{N} = (\mathbb{N}, 0, S, +, \times)$:

- $\mathbb{N}$ is the set of natural numbers,
- $0 \in \mathbb{N}$ is zero (the least natural number),
- $S : \mathbb{N} \to \mathbb{N}$ is the *successor function* $S(n) = n + 1$,
- $+$ and $\times$ are addition and multiplication of natural numbers in the usual sense.

It turns out the things are now *way more* non-trivial.

- For *any* infinite $\mathcal{A}$, there is no Σ that "determines" $\mathcal{A}$
  (Interested students can check the *Lövenheim- Skolem theorem*)
- Even worse, there is no desired Σ for $\mathcal{N}$ $\leftarrow$ our goal!

(One of) our last goals—the formal statement

Now we can formally state one of our eventual goals.

Here, **Q** is a set of sentences called **Robinson arithmetic**: For now, just assume it is a certain set of sentences that describes an elementary aspect of $\mathcal{N}$.

Theorem (corollary of Gödel's first incompleteness theorem)

There is no recursive complete extension of **Q** *modeled by* $\mathcal{N}$.