Logic and Computation II

Part 4. Formal arithmetic and Gödel's incompleteness theorems

Kazuyuki Tanaka

BIMSA

March 14, 2023

北京雁栖湖应用数学研究院
YANQI LAKE BEIJING INSTITUTE OF MATHEMATICAL SCIENCES AND APPLICATIONS

Logic and Computation II

- **Part 4. Formal arithmetic and Gödel's incompleteness theorems**
- **Part 5. Automata on infinite objects**
- **Part 6. Recursion-theoretic hierarchies**
- **Part 7. Admissible ordinals and second order arithmetic**

Part 4. Schedule

- Mar. 7, (1) First-order logic
- Mar. 9, (2) Arithmetical formulas
- Mar.14, (3) Gödel's first incompleteness theorem
- Mar.16, (4) Gödel's second incompleteness theorem
- Mar.21, (5) Second-order logic
- Mar.23, (6) Analytical formulas

Today's topics

Recap: Peano Arithmetic

Peano arithmetic is a first-order theory in the language of ordered rings $\mathcal{L}_{\mathrm{OR}} = \{+, \cdot, 0, 1, <\}$.

Definition

Peano arithmetic (PA) consists of the following axioms.

Successor:	$\neg(x+1=0)$,	$x+1=y+1 \to x=y$.
Addition:	$x+0=x$,	$x+(y+1)=(x+y)+1$.
Multiplication:	$x \cdot 0 = 0$,	$x \cdot (y+1) = x \cdot y + x$.
Inequality	$\neg(x<0)$,	$x<y+1 \leftrightarrow x<y \vee x=y$.
Induction:	$\varphi(0) \wedge \forall x(\varphi(x) \to \varphi(x+1)) \to \forall x \varphi(x)$.	

- Induction is not a single formula, but an axiom schema that collects the formulas for all the $\varphi(x)$ in $\mathcal{L}_{\mathrm{OR}}$. Note that $\varphi(x)$ may include free variables other than x.

Arithmetical Hierarchy

- We inductively define hierarchical classes of formulas Σ_i and Π_i $(i \in \mathbb{N})$.

Definition

- The **bounded** formulas are constructed from atomic formulas by using propositional connectives and bounded quantifiers $\forall x < t$ and $\exists x < t$, where $\forall x < t$ and $\exists x < t$ are abbreviations for $\forall x(x < t \to \cdots)$ and $\exists x(x < t \land \cdots)$, respectively, and t is a term that does not include x. A bounded formula is also called a Σ_0 $(=\Pi_0)$ formula.
- For any $i, k \in \mathbb{N}$:
  - if φ is a Σ_i formula, $\forall x_1 \cdots \forall x_k \varphi$ is a Π_{i+1} formula,
  - if φ is a Π_i formula, $\exists x_1 \cdots \exists x_k \varphi$ is a Σ_{i+1} formula.

- If a Π_i formula is equivalent to some Σ_i formula or a Σ_i formula equivalent to some Π_i formula, such a formula is called a Δ_i formula.

Let us define subsystems of Peano arithmetic PA by restricting its induction axiom.

Definition

Let Γ be a class of formulas in $\mathcal{L}_{\mathrm{OR}}$. By I$\Gamma$, we denote a subsystem of PA obtained by restricting ($\varphi(x)$ of) induction to the class Γ.

- The main subsystems of PA are $\mathrm{I}\Sigma_1 \supset \mathrm{I}\Sigma_0 \supset \mathrm{IOpen}$, where Open is the set of formulas without quantifiers.

Another system weaker than IOpen is the system Q defined by R. Robinson.

Definition

Robinson's system Q is obtained from PA by removing the axioms of inequality and induction, and instead adding the following axiom:

Predecessor: $\forall x(x \neq 0 \to \exists y(y+1=x))$.

So, it is a theory in the language of ring $\mathcal{L}_{\mathrm{R}} = \{+, \cdot, 0, 1\}$.

Let $\mathrm{Q}_<$ be the system Q plus the definition of the inequality symbol.

Lemma

In IOpen, all axioms of **theory of discrete ordered semirings** PA^- can be proved.

(1) Semiring axiom (excluding the existence of additive inverses from the commutative ring axiom).

(2) difference axiom $x < y \to \exists z(z + (x+1) = y)$.

(3) 0 as the minimum element in linear order and discrete $(0 < x \leftrightarrow 1 \leq x)$.

(4) Order preservation $x < y \to x + z < y + z \land (x \cdot z < y \cdot z \lor z = 0)$.

Corollary

$\mathsf{Q}_< \subset \mathsf{PA}^- \subset \mathrm{IOpen} \subset \mathrm{I}\Sigma_0 \subset \mathrm{I}\Sigma_1 \subset \mathsf{PA}$.

- In $\mathsf{Q}_<$, an atomic formula $s = t$ or $s < t$ without variables can be proved if true, and its negation can be proved if false (by meta-induction on the composition of terms).
- Furthermore, a bounded formula without free variables can be proved/disproved in $\mathsf{Q}_<$ if it is true/false.
- A system is said to be Σ_1-**complete** if it proves all true Σ_1 sentences.

Theorem (Σ_1-completeness of $\mathsf{Q}_<$)

$\mathsf{Q}_<$ proves all true Σ_1 sentences.

Proof

- If a Σ_1 sentence $\exists x_1 \exists x_2 \ldots \exists x_k \varphi(x_1, x_2, \ldots, x_k)$ is true, there exist concrete numbers $n_1, n_2, \ldots, n_k$ such that $\varphi(\overline{n_1}, \overline{n_2}, \ldots, \overline{n_k})$ holds, where $\overline{n} = \overbrace{1 + \cdots + 1}^{n \text{ times}}$ and $\overline{n} = 0$.
- Since $\varphi(\overline{n_1}, \overline{n_2}, \ldots, \overline{n_k})$ is a bounded formula, it is provable if it is true. From the rule of first-order logic, $\exists x_1 \exists x_2 \ldots \exists x_k \varphi(x_1, x_2, \ldots, x_k)$ is also provable. $\square$

- All the arithmetic systems we will discuss are extensions of $\mathsf{Q}_<$, and thus Σ_1-complete.
- Another condition for a theory to induce the first incompleteness theorem is **1-consistency**, also known as Σ_1-**soundness**. A theory is said to be Σ_n-**sound** if all provable Σ_n statements are true.
- Gödel's original condition, called ω-**consistency** is strictly stronger than Σ_1-soundness.

- We first look at the first incompleteness theorem from the viewpoint of computability theory. Then, we will reexamine the proof more syntactically.
- Recall that $X \subseteq \mathbb{N}^n$ is called **CE** (computably enumerable) if it is the domain (or range) of some partial recursive function. Then, from the lemma below, any CE relation $R(\vec{x})$ can be expressed by $\exists y S(\vec{x}, y)$ for some primitive recursive relation S.

Recall, Lemma in Lecture-01-05 of this course

For the relation $R \subset \mathbb{N}^n$, the following conditions are equivalent.

(1) R is recursively enumerable (CE).

(6) There exists a primitive recursive relation S such that
$$R(x_1, \cdots, x_n) \Leftrightarrow \exists y S(x_1, \cdots, x_n, y).$$

(7) There exists a recursive relation S such that
$$R(x_1, \cdots, x_n) \Leftrightarrow \exists y S(x_1, \cdots, x_n, y).$$

Definition

Let $\mathfrak{N} = (\mathbb{N}, +, \cdot, 0, 1, <)$ be a standard model of PA.

- A set $A \subseteq \mathbb{N}^l$ is said to be Σ_i if there exists a Σ_i formula $\varphi(x_1, \dots, x_l)$ satisfying

$$(m_1, \dots, m_l) \in A \Leftrightarrow \mathfrak{N} \models \varphi(\overline{m_1}, \dots, \overline{m_l}).$$

- Here, $\overline{m}$ is a term expressing number m, that is, $\overline{m} = \overbrace{(1 + 1 + \cdots + 1)}^{m}(m > 0)$, $\overline{0} = 0$.
- Similarly, Π_i sets can be defined by Π_i formulas.
- A set that is both Σ_i and Π_i is called Δ_i.

- By Lemma (2) later, we will show that the Σ_1 sets are the CE sets.

Lemma (1)

The graph $\{(\vec{x}, y) : f(\vec{x}) = y\}$ of a primitive recursive function f is a Δ_1 set.

Proof

- By induction on the construction of primitive recursive functions. The main part is to treat the definition by primitive recursion.
- For simplicity, we omit parameter variables $x_1, \ldots, x_l$, and consider the definition of a unary function f from a constant c and binary function h as follows:

$$f(0) = c, \quad f(y+1) = h(y, f(y)).$$

- From the induction hypothesis, h can be expressed in both Σ_1 and Π_1 formulas.
- First, let $\gamma(x, m, n)$ be a Σ_0 formula expressing "$m(x+1)+1$ is a divisor of n", that is, $\exists d < n \ (m(x+1)+1) \cdot d = n$. Then, for any finite set A (with $\max A < u$), there exist m, n such that $\forall x < u(x \in A \Leftrightarrow \gamma(x, m, n))$.
- In fact, assume $(u-1)! \mid m$. Then, $(m(i+1)+1)$ and $(m(j+1)+1)$ are mutually prime for any $i < j < u$. Thus, $n = \Pi_{i \in A}(m(i+1)+1)$ works.

- Now, we will define a Σ_0 formula $\delta(u, m, n)$ such that

$$\delta(\langle u_1, u_2 \rangle, m, n) \Leftrightarrow \forall y < u_1 \exists z < u_2 \ f(y) = z.$$

- The formula $\delta(u, m, n)$ is formally constructed as follows: for any $u = \langle u_1, u_2 \rangle$,

$$\delta(u, m, n) \equiv \forall y < u_1 \exists z < u_2 \ \gamma(\langle y, z \rangle, m, n) \land \forall z < u_2 (\gamma(\langle 0, z \rangle, m, n) \leftrightarrow z = c)$$
$$\land \forall y < u_1 - 1 \forall z < u_2 (\gamma(\langle y+1, z \rangle, m, n) \leftrightarrow \exists z' < u_2 (z = h(y, z') \land \gamma(\langle y, z' \rangle, m, n))).$$

- Then $\forall u_1 \exists u_2 \exists m \exists n \delta(\langle u_1, u_2 \rangle, m, n)$ holds. Thus, we obtain

$$\begin{aligned} f(y) = z &\Leftrightarrow \exists u \exists m \exists n (u_1 = y + 1 \land \delta(u, m, n) \land \gamma(\langle y, z \rangle, m, n)) \\ &\Leftrightarrow \forall u \forall m \forall n (u_1 = y + 1 \land \delta(u, m, n) \rightarrow \gamma(\langle y, z \rangle, m, n)). \end{aligned}$$

- That is, $f(y) = z$ is a Δ_1 set. □

- As we saw in the revisited lemma on Slides p.9, any CE relation $R(\vec{x})$ can be expressed by $\exists y S(\vec{x}, y)$ for some primitive recursive relation S.
- By the above lemma, the primitive recursive relation S can be expressed by a Σ_1 formula, and $\exists y S(\vec{x}, y)$ is still Σ_1. Thus, any CE relation can be expressed by a Σ_1 formula.
- Therefore, we have the following.

Lemma (2)

The CE sets are exactly the same as the Σ_1 sets. Hence, the computable (recursive) sets are exactly the same as the Δ_1 sets.

Then, the following two formal representation theorems hold.

Theorem ((Weak) Representation Theorem for CE sets)

Suppose that a theory T is Σ_1-complete and 1-consistent. Then, for any CE set C, there exists a Σ_1 formula $\varphi(x)$ such that for any n,

$$n \in C \quad \Leftrightarrow \quad T \vdash \varphi(\overline{n}).$$

Proof.

- From the Lemma (2), for any CE set C, there exists a Σ_1 formula $\varphi(x)$ such that $n \in C \Leftrightarrow \mathfrak{N} \models \varphi(\overline{n})$.
- Since T is Σ_1-complete, $\mathfrak{N} \models \varphi(\overline{n}) \Rightarrow T \vdash \varphi(\overline{n})$.
- Also because T is 1-consistent, $T \vdash \varphi(\overline{n}) \Rightarrow \mathfrak{N} \models \varphi(\overline{n})$.

□

Theorem ((Strong) Representation Theorem for Computable Sets)

Assume a theory T is Σ_1-complete. For any computable set C, there exists a Σ_1 formula $\varphi(x)$ such that

$$n \in C \Rightarrow T \vdash \varphi(\overline{n}), \quad n \notin C \Rightarrow T \vdash \neg\varphi(\overline{n}).$$

Proof.

- For a computable set C, from the Lemma (2) there exist Σ_0 formulas $\theta_1(x, y), \theta_2(x, y)$ such that

$$n \in C \Leftrightarrow \mathfrak{N} \models \exists y \theta_1(\overline{n}, y), \quad n \notin C \Leftrightarrow \mathfrak{N} \models \exists y \theta_2(\overline{n}, y).$$

  Now, let $\varphi(x)$ be a Σ_1 formula $\exists y(\theta_1(\overline{n}, y) \wedge \forall z \leq y \neg\theta_2(\overline{n}, z))$. By the Σ_1-completeness of T, $n \in C \Rightarrow T \vdash \varphi(\overline{n})$.
- To show $n \notin C \Rightarrow T \vdash \neg\varphi(\overline{n})$, let $n \notin C$.
  Then, since $\mathfrak{N} \models \exists y \theta_2(\overline{n}, y)$, some m exists and $\mathfrak{N} \models \theta_2(\overline{n}, \overline{m})$. From the Σ_1 completeness of T, $T \vdash \theta_2(\overline{n}, \overline{m})$.
  Also, since $\mathfrak{N} \not\models \exists y \theta_1(\overline{n}, y)$, for all l, $\mathfrak{N} \models \neg\theta_1(\overline{n}, \overline{l})$, i.e., $T \vdash \neg\theta_1(\overline{n}, \overline{l})$.
  Therefore, if $\theta_1(\overline{n}, a)$ in some model of T, then a is not a standard natural number l.
  Thus, $T \vdash \forall y(\theta_1(\overline{n}, y) \to \exists z \leq y\ \theta_2(\overline{n}, z))$, that is, $T \vdash \neg\varphi(\overline{n})$. □

- To derive the incompleteness theorem, we need one more condition on a formal system, that is, the set of axioms is CE.
- Without this condition, for example, if we take all true arithmetic formulas as axioms, we would have a complete theory, but it would not be a formal system.
- From the following theorem, the CE set of axioms can be also expressed as a primitive recursive set.

Theorem (Craig's lemma)

For any CE theory T, there exists an equivalent (proving the same theorem) primitive recursive theory T'.

Proof. Let T be a CE theory, defined by Σ_1 formula $\varphi(x) \equiv \exists y \theta(x, y)$ (θ is Σ_0).
That is, $\sigma \in T \Leftrightarrow \mathfrak{N} \models \varphi(\overline{\ulcorner \sigma \urcorner})$. $\ulcorner \sigma \urcorner$ is the Gödel number of a sentence σ.
Then, we define a primitive recursive theory T' as follows:

$$T' = \{\overbrace{\sigma \wedge \sigma \wedge \cdots \wedge \sigma}^{n+1 \text{ copies}} : \theta(\overline{\ulcorner \sigma \urcorner}, \overline{n})\}.$$

Then, T and T' are equivalent, since $\vdash \sigma \leftrightarrow \sigma \wedge \sigma \wedge \cdots \wedge \sigma$.

In the proof above, the definition of T' is not Σ_0 since it includes the Gödel numbers, etc. The following can be shown about the CE theory.

Theorem

For any CE theory T, the set of its theorems $\{\ulcorner\sigma\urcorner : T \vdash \sigma\}$ is also CE.

Proof

- Recall that a proof in a formal system of first-order logic is a finite sequence of formulas, each formula being either a logical axiom, an equality axiom, or a mathematical axiom of a theory T, or obtained from previous formulas by applying MP or a quantification rule.
- From the Craig's Lemma, a CE theory T can be transformed into a primitive recursive theory. Thus, it is also a primitive recursive relation that (the Gödel number of) a finite sequence of formulas is a proof of T.
- The set of theorems of T is CE. Because a sentence σ is a theorem of T iff there exists a proof (i.e., a sequence that satisfies the primitive recursive relation) such that σ is the last formula of the proof. □

The halting problem K is CE, but its complement $\mathbb{N} - \mathrm{K}$ is not (part 1 of this course). Gödel's first incompleteness theorem easily follows from this fact.

Theorem (Gödel's first incompleteness theorem)

Let T be a Σ_1-complete and 1-consistent CE theory. Then T is incomplete, that is, there is a sentence that cannot be proved or disproved.

Proof.

- Suppose K is CE but not co-CE. By the weak representation theorem for CE sets, there exists a formula $\varphi(x)$ such that

$$n \in \mathrm{K} \Leftrightarrow T \vdash \varphi(\overline{n}).$$

- On the other hand, since $\mathbb{N} - \mathrm{K}$ is not a CE, there exists some d such that

$$d \in \mathbb{N} - \mathrm{K} \not\Leftrightarrow T \vdash \neg\varphi(\overline{d}).$$

  Thus, ($d \in \mathrm{K}$ and $T \vdash \neg\varphi(\overline{d})$) or ($d \notin \mathrm{K}$ and $T \nvdash \neg\varphi(\overline{d})$).
  - In the former case, since $d \in \mathrm{K}$ implies $T \vdash \varphi(\overline{d})$, T is inconsistent, contradicting with the 1-consistency assumption.
  - In the latter case, T is incomplete because $\varphi(\overline{d})$ cannot be proved or disproved.

Homework

(1) In a Σ_1 complete theory T, show that 1-consistency (Σ_1-soundness) of T is equivalent to the following: for any Σ_0 formula $\varphi(x)$, if $\varphi(\overline{n})$ is provable in T for all n, then $\exists x \neg\varphi(x)$ is not provable in T.

(2) Let A, B be two disjoint CE sets. Assume a theory T is Σ_1-complete. Show that there exists a Σ_1 formula $\psi(x)$ such that

$$n \in A \Rightarrow T \vdash \psi(\overline{n}), \quad n \in B \Rightarrow T \vdash \neg\psi(\overline{n}).$$

From this, deduce that $\{\ulcorner\sigma\urcorner : T \vdash \sigma\}$ and $\{\ulcorner\sigma\urcorner : T \vdash \neg\sigma\}$ are computably inseparable. (See Part 1-6, Slide p.25.) In particular, $\{\ulcorner\sigma\urcorner : T \vdash \sigma\}$ is not computable.

From now, we will reconsider the proof of the first incompleteness theorem more rigorously and constructively, which leads us to the second incompleteness theorem.
By the previous lemma, we know that the graph of a primitive recursive (p.r.) function f is a Δ_1 set. In fact, we can prove the following.

Lemma (Strong Representation for primitive recursive functions)

For any p.r. function, there is a Δ_1 formula $\varphi(x, y)$ which expresses its graph, and moreover in $\mathrm{I}\Sigma_1$, it is provable that $\forall x \exists ! y \varphi(x, y)$. Here, $\exists !$ means "there is a unique".

Proof.
$\forall x \exists ! y \varphi(x, y)$ is based on the fact that $\varphi(x, y)$ represents a p.r. function, which can be shown by induction on $\mathrm{I}\Sigma_1$, since $\varphi(x, y)$ is a Δ_1 formula. □

- The constructive proof of the first incompleteness theorem utilizes the diagonalization lemma.
- To state the lemma, we need the following fact: $\mathrm{I}\Sigma_1$ is a conservative extension even if we add the symbols of the p.r. function and its definable formulas.
- In other words, (in $\mathrm{I}\Sigma_1$) a Σ_1 formula containing a p.r. functions is rewritten as an equivalent Σ_1 formula without p.r. functions by replacing the functions with the corresponding Δ_1 formula.

Lemma (Diagonalization lemma)

Let T be any extension of $\mathrm{I}\Sigma_1$. For any formula $\psi(x)$ in which x is the only free variable, there exists a sentence σ such that $T \vdash$ "$\sigma \leftrightarrow \psi(\overline{\ulcorner \sigma \urcorner})$" .

Proof.

- A formula with only x as a free variable is primitively recursively enumerated as $\varphi_0(x), \varphi_1(x), \ldots$, and then $f(n) = \ulcorner \varphi_n(\overline{n}) \urcorner$ is also a p.r. function. By the strong representation theorem for p.r. functions, there exists a Σ_1 formula χ such that

$$f(m) = n \Rightarrow T \vdash \chi(\overline{m}, \overline{n}) \land \forall x \exists ! y \chi(x, y).$$

- Considering the formula $\exists y (\chi(x, y) \land \psi(y))$, since it only has free variable x, it is $\varphi_k(x)$ for some k.
- Let $\sigma \equiv \varphi_k(\overline{k})$ for this k. Then, $f(k) = \ulcorner \sigma \urcorner$, so $T \vdash \chi(\overline{k}, \overline{\ulcorner \sigma \urcorner})$.
- Thus, in T,

$$\psi(\overline{\ulcorner \sigma \urcorner}) \to \exists y (\chi(\overline{k}, y) \land \psi(y)) \ (\equiv \varphi_k(\overline{k}) \equiv \sigma)$$

- On the other hand, since $T \vdash \forall x \exists ! y \chi(x, y)$, in T,

$$\neg \psi(\overline{\ulcorner \sigma \urcorner}) \to \neg \exists y (\chi(\overline{k}, y) \land \psi(y)) \ (\equiv \neg \sigma).$$

- Therefore, σ is the fixed point of ψ ($T \vdash \sigma \leftrightarrow \psi(\overline{\ulcorner \sigma \urcorner})$). □

Based on Craig's lemma, a CE theory is primitive recursively axiomatizable. Then, "a finite sequence of formulas P is a proof in T" can be defined in a primitive recursive way.

Definition

- Let T be a CE theory and its p.r. counterpart T'.
- A proof in T' is a finite sequence of formulas where each formula is either a logical axiom, an equality axiom, or an axiom of T', or obtained by applying MP or quantification rules from formulas.
- The formula that appears at the end of the proof is the theorem of T.
- Now, we define the predicate Proof_T as follows.

$$\mathrm{Proof}_T(\ulcorner P \urcorner, \ulcorner \sigma \urcorner) \Leftrightarrow P \text{ is a proof of formula } \sigma \text{ in } T'.$$

- By Proof_T, we also denote a Δ_1 formula expressing the above Proof_T in $\mathsf{I}\Sigma_1$. A Σ_1 formula Bew_T is defined as $\mathrm{Bew}_T(x) \equiv \exists y\ \mathrm{Proof}_T(y, x)$.

The formula $\mathrm{Bew}_T(x)$ expresses that "x is the Gödel number of a theorem of T". "Bew" stands for the German beweisbar (provable).

Alternative proof: the first incompleteness

We give another proof for the first incompleteness theorem (with the additional assumption that a theory T includes $\mathsf{I}\Sigma_1$).

Proof.

- By the diagonalization lemma, $\neg\mathrm{Bew}_T(x)$ has a fixed point, that is, there exists σ such that $T \vdash \sigma \leftrightarrow \neg\mathrm{Bew}_T(\overline{\ulcorner\sigma\urcorner})$.
- We will show this σ is neither provable nor disprovable in T as follows.
- Let $T \vdash \sigma$. Then $\mathrm{Bew}_T(\overline{\ulcorner\sigma\urcorner})$ is true. Hence $T \vdash \mathrm{Bew}_T(\overline{\ulcorner\sigma\urcorner})$ from Σ_1 completeness. Since σ is the fixed point of $\neg\mathrm{Bew}_T(x)$, we have $T \vdash \neg\sigma$, which means that T is inconsistent.
- On the other hand, if $T \vdash \neg\sigma$, $T \vdash \mathrm{Bew}_T(\overline{\ulcorner\sigma\urcorner})$ because σ is a fixed point. Here, using the 1-consistency of T, $\mathrm{Bew}_T(\overline{\ulcorner\sigma\urcorner})$ is true, and so $T \vdash \sigma$, which is a contradiction. □

The sentence σ thus constructed "asserts its own unprovability" because "$\sigma \Leftrightarrow T \nvdash \sigma$" holds. This σ is called the **Gödel sentence** of T.

Summary

Theorem (Gödel's first incompleteness theorem)

Any 1-consistent CE extension of $\mathsf{I}\Sigma_1$ is incomplete.

Further readings

- Theory of Computation, D.C. Kozen, Springer 2006.
- Mathematical Logic. H.-D. Ebbinghaus, J. Flum, W. Thomas, Graduate Texts in Mathematics 291, Springer 2021.