Logic and Computation I
Part 3a. Formal Arithmetic

Kazuyuki Tanaka

BIMSA

December 3, 2024

北京雁栖湖应用数学研究院
Beijing Institute of Mathematical Sciences and Applications

Logic and Computation I

- **Part 1. Introduction to Theory of Computation**
- **Part 2. Propositional Logic and Computational Complexity**
- **Part 3. First Order Logic and Decision Problems**
- **Part 3a. Formal Arithmetic**

Part 3a. Schedule (subject to change)

- Nov.21, (6) Presburger arithmetic
- Nov.26, (7) Peano arithmetic
- Nov.28, (8) Gödel's first incompleteness theorem
- **Dec. 3, (9) Gödel's second incompleteness theorem**
- Dec. 5, (10) Second order logic
- Dec.10, (11) Second order arithmetic

Lemma 3.53 (Formal representation for primitive recursive functions)

For any primitive recursive function f, there is a Δ_1 formula $\chi(x, y)$ such that

$$f(m) = n \Rightarrow \mathsf{I}\Sigma_1 \vdash \chi(\overline{m}, \overline{n}) \quad \text{and} \quad \mathsf{I}\Sigma_1 \vdash \forall x \exists! y \chi(x, y).$$

Then, $\mathsf{I}\Sigma_1 + \forall x \chi(x, f(x))$ is conservative over $\mathsf{I}\Sigma_1$.

Lemma 3.54 (Diagonalization lemma)

For any formula $\psi(x)$ with a unique free variable x, there exists a sentence σ such that $\mathsf{I}\Sigma_1 \vdash \sigma \leftrightarrow \psi(\overline{\ulcorner \sigma \urcorner})$.

Definition 3.55 (Provability predicate Bew)

For a CE theory T, we define a prim. rec. relation $\mathrm{Proof}_T(\ulcorner P \urcorner, \ulcorner \sigma \urcorner)$ to express "P is a proof of formula σ in T". By Proof_T, we also denote a Δ_1 formula expressing Proof_T in $\mathsf{I}\Sigma_1$. A Σ_1 formula Bew_T is defined as $\exists y\, \mathrm{Proof}_T(y, x)$.

$\mathrm{Bew}_T(x)$ expresses that "x is the Gödel number of a theorem of T".

Gödel's first incompleteness theorem

Any 1-consistent CE theory T including $\mathsf{I}\Sigma_1$ is incomplete.

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 a 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.

Using the exercise problem in the previous lecture, the assumption of Gödel's theorem can be weakened from 1-consistency to consistency.

Gödel-Rosser's incompleteness theorem

Any **consistent** CE theory T including $\mathsf{I}\Sigma_1$ is incomplete.

Proof.

- Let $A = \{\ulcorner\sigma\urcorner : T \vdash \sigma\}$, $B = \{\ulcorner\sigma\urcorner : T \vdash \neg\sigma\}$. If T is consitent CE theory, then A, B are disjoint CE sets.
- Similarly to the proof of the strong representation theorem (3.49) for computable sets, costruct a formula $\psi(x)$ such that $A \subset \{n : T \vdash \psi(n)\}$ and $B \subset \{n : T \vdash \neg\psi(n)\}$.
- By the diagonalization lemma (3.54), we have a sentence σ such that $T \vdash (\sigma \leftrightarrow \neg\psi(\ulcorner\sigma\urcorner))$, and can prove that $\ulcorner\sigma\urcorner \notin A \cup B$.

Two applications of the first incomp. theorem

The next theorem is also a very important corollary of the argument of the first incompleteness theorem. Note that T in the diagonalization lemma does not need be a CE theory. So, letting T be $\mathrm{Th}(\mathfrak{N})$ (the set of sentences true in $\mathfrak{N}$), we have

Theorem 3.56 (Tarski's Truth Indefinability)

For any sentence σ, there is no formula $\psi(x)$ such that

$$\mathfrak{N} \models \sigma \leftrightarrow \psi(\overline{\ulcorner \sigma \urcorner}).$$

In other words, $\{\ulcorner \sigma \urcorner : \mathfrak{N} \models \sigma\}$ is not arithmetically definable.

Proof. Consider a fixed point σ for $\neg\psi(x)$.

The following theorem was due to Church. Turing also obtained a similar result by expressing the halting problem as a satisfaction problem of first-order logic.

Theorem 3.57 (Undecidability of first-order logic)

$\{\ulcorner\sigma\urcorner : \ \sigma$ is a valid sentence in the language $\mathcal{L}_{\mathrm{OR}}\}$ is not computable. Therefore, the satisfiability of first order logic is not decidable.

Proof.

- First note that $\mathsf{I}\Sigma_1$ is finitely axiomatizable, because the Σ_1-induction schema can be expressed as a single induction axiom for a universal Σ_1-formula (a universal CE set). Or, instead of $\mathsf{I}\Sigma_1$, you may take $\mathsf{Q}_<$ or any other finitely axiomatized theory for which the first incompleteness theorem can be shown.
- Let ξ be a sentence obtained by connecting all the axioms of $\mathsf{I}\Sigma_1$ by $\wedge$. From the deduction theorem, $\mathsf{I}\Sigma_1 \vdash \sigma \Leftrightarrow \vdash \xi \to \sigma$. If $\{\ulcorner\sigma\urcorner : \vdash \sigma\}$ were computable, $\{\ulcorner\sigma\urcorner : \vdash \xi \to \sigma\} = \{\ulcorner\sigma\urcorner : \mathsf{I}\Sigma_1 \vdash \sigma\}$ would also be computable, which leads to contradiction by diagonalization (as in the argument on p.5). So, by the completeness theorem, the validity of a sentence is not decidable.
- Since the satisfiability of a sentence σ can be expressed as $\not\models \neg\sigma$, it is also not computable.

Intoducing the second incompleteness theorem

- A version of the first incompleteness theorem says that a consitent CE theory T including $\mathsf{I}\Sigma_1$ (indeed $\mathsf{Q}_<$ is enough) neither prove (nor disprove) the Gödel sentence.
- A main part of the second incompleteness theorem says that a CE theory T including $\mathsf{I}\Sigma_1$ proves that the consistency of T implies the Gödel sentence (equivalently, its unprovability).
- Then, we obtain the second incompleteness theorem that a consistent T does not prove its consistency, since if it did then it would also prove the Gödel sentence, which contradicts with the first theorem.
- Thus, the main part of the proof of the second theorem is to formalize the proof of the first theorem in the system T.
- Although this requires extremely elaborate arguments, the main points are summarized as the three properties of the derivability predicate $\mathrm{Bew}_T(x)$ as shown in the next slide.

Lemma 3.58 (Hilbert-Bernays-Löb's derivability condition)

Let T be a consistent CE theory containing IΣ_1. For any φ, ψ,

D1. $T \vdash \varphi \Rightarrow T \vdash \mathrm{Bew}_T(\overline{\ulcorner \varphi \urcorner})$.

D2. $T \vdash \mathrm{Bew}_T(\overline{\ulcorner \varphi \urcorner}) \land \mathrm{Bew}_T(\overline{\ulcorner \varphi \to \psi \urcorner}) \to \mathrm{Bew}_T(\overline{\ulcorner \psi \urcorner})$.

D3. $T \vdash \mathrm{Bew}_T(\overline{\ulcorner \varphi \urcorner}) \to \mathrm{Bew}_T(\overline{\ulcorner \mathrm{Bew}_T(\overline{\ulcorner \varphi \urcorner}) \urcorner})$.

Proof

- D1 is obtained from the Σ_1 completeness of T, since $\mathrm{Bew}_T(\overline{\ulcorner \varphi \urcorner})$ is a Σ_1 formula.
- For D2, it is clear that the proof of ψ is obtained by applying MP to the proof of φ and the proof of $\varphi \to \psi$.
- D3 formalizes D1 in T. This is the most difficult, since we can not find a simple machinery to transform a proof of φ in T to a proof of $\mathrm{Bew}_T(\overline{\ulcorner \varphi \urcorner})$. We will explain an idea of this machinery in the next slide.

- First, we prove that, for any primitive recursive function f,

$$T \vdash f(x_1, \dots, x_k) = y \to \mathrm{Bew}_T(\overline{\ulcorner f(\dot{x_1}, \dots, \dot{x_k}) = \dot{y} \urcorner}).$$

  Here, the function $\dot{x}$ is a primitive recursive function from a number n to the Gödel number of its numeral $\ulcorner \bar{n} \urcorner$.
- The above formula can be proved by meta-induction on the construction of the primitive recursive function f.
- Now, assume $\mathrm{Bew}_T(\overline{\ulcorner \varphi \urcorner})$. Then, there is a numeral c that satisfies $\mathrm{Proof}_T(c, \overline{\ulcorner \varphi \urcorner})$. So, substituing (the numeral of the Gödel number of) this formula into $\mathrm{Bew}_T(x)$, we finally obtain $\mathrm{Bew}_T(\overline{\ulcorner \mathrm{Bew}_T(\overline{\ulcorner \varphi \urcorner}) \urcorner})$ by a simple computation.
- For more details, please refer to my book[1] or other.
- Another proof will be given later.

□

[1] https://www.shokabo.co.jp/mybooks/ISBN978-4-7853-1575-7.htm

In the following, let π_G denote a Gödel sentence such that

$$T \vdash \pi_G \leftrightarrow \neg \mathrm{Bew}_T(\overline{\ulcorner \pi_G \urcorner}).$$

By $\mathrm{Con}(T)$, we denote the sentence meaning "T is consistent", formally defined as

$$\mathrm{Con}(T) \equiv \neg \mathrm{Bew}_T(\overline{\ulcorner 0 = 1 \urcorner}).$$

Then we have the following.

Lemma 3.59

$T \vdash \mathrm{Con}(T) \leftrightarrow \pi_G$.

Proof. • To show $\pi_G \to \mathrm{Con}(T)$. $T \vdash 0 = 1 \to \pi_G$, so by D1 and D2,

$$T \vdash \mathrm{Bew}_T(\overline{\ulcorner 0 = 1 \urcorner}) \to \mathrm{Bew}_T(\overline{\ulcorner \pi_G \urcorner}).$$

Taking the contraposition, we get $T \vdash \pi_G \to \mathrm{Con}(T)$.

Proof. • To show $\mathrm{Con}(T) \to \pi_G$.
First, from $T \vdash \pi_G \leftrightarrow \neg \mathrm{Bew}_T(\overline{\ulcorner \pi_G \urcorner})$ and D1,

$$T \vdash \mathrm{Bew}_T(\overline{\ulcorner \mathrm{Bew}_T(\overline{\ulcorner \pi_G \urcorner}) \to \neg \pi_G \urcorner}).$$

Using D2,

$$T \vdash \mathrm{Bew}_T(\overline{\ulcorner \mathrm{Bew}_T(\overline{\ulcorner \pi_G \urcorner}) \urcorner}) \to \mathrm{Bew}_T(\overline{\ulcorner \neg \pi_G \urcorner}).$$

By D3, $T \vdash \mathrm{Bew}_T(\overline{\ulcorner \pi_G \urcorner}) \to \mathrm{Bew}_T(\overline{\ulcorner \mathrm{Bew}_T(\overline{\ulcorner \pi_G \urcorner}) \urcorner})$, so

$$T \vdash \mathrm{Bew}_T(\overline{\ulcorner \pi_G \urcorner}) \to \mathrm{Bew}_T(\overline{\ulcorner \neg \pi_G \urcorner}).$$

Using $T \vdash \pi_G \to (\neg \pi_G \to 0 = 1)$ and D2, from above

$$T \vdash \mathrm{Bew}_T(\overline{\ulcorner \pi_G \urcorner}) \to \mathrm{Bew}_T(\overline{\ulcorner 0 = 1 \urcorner})$$

Taking the contraposition,

$$T \vdash \neg \mathrm{Bew}_T(\overline{\ulcorner 0 = 1 \urcorner}) \to \neg \mathrm{Bew}_T(\overline{\ulcorner \pi_G \urcorner}),$$

That is, $T \vdash \mathrm{Con}(T) \to \pi_G$. □

Theorem 3.60 (Gödel's second incompleteness theorem)

Let T be a consistent CE theory, which contains $\mathrm{I}\Sigma_1$. Then $\mathrm{Con}(T)$ cannot be proved in T.

Proof

By the proof of the first incompleteness theorem, $T \nvdash \pi_G$.
By the above lemma, $T \vdash \mathrm{Con}(T) \leftrightarrow \pi_G$, so $T \nvdash \mathrm{Con}(T)$. □

Remark

In mathematical logic, the second incompleteness theorem is often used to separate two axiomatic theories by showing the consistency of one over the other. E.g. $\mathrm{I}\Sigma_1$ is a proper subsystem of PA, since the consistency of the former can be proved in the latter.

Exercise

(1) Show that there is a consistent theory T that proves its own inconsistency $\neg\mathrm{Con}(T)$.

(2) Let $\mathrm{Bew}^{\#}_T(x) \equiv (\mathrm{Bew}_T(x) \land x \neq \overline{\ulcorner 0 = 1 \urcorner})$. For any true proposition σ,

$$\mathrm{Bew}^{\#}_T(\overline{\ulcorner \sigma \urcorner}) \leftrightarrow \mathrm{Bew}_T(\overline{\ulcorner \sigma \urcorner})$$

and

$$T \vdash \neg\mathrm{Bew}^{\#}_T(\overline{\ulcorner 0 = 1 \urcorner}).$$

Does it contradict with the second incompleteness theorem?

Model-theoretic arguments due to Kikuchi-Tanaka

- In T, we can prove a countable version of the completeness theorem of first-order logic. A countable model M can be treated as its coded diagram, i.e., the set of the Gödel numbers of $\mathcal{L}_M$-sentences true in M. The arithmetized completeness theorem says that if T' is consistent then there exists (a formula expressing the diagram of) a model of T'.
- To show D3, we work in $T + \mathrm{Bew}_T(\overline{\varphi})$. For any model M of T, $\mathrm{Bew}_T(\overline{\varphi})$ holds in M by Σ_1 completeness. Hence, by Gödel completeness, we have $\mathrm{Bew}_T(\overline{\ulcorner \mathrm{Bew}_T(\overline{\ulcorner \pi_G \urcorner}) \urcorner})$.
- We can directly prove $\mathrm{Con}(T) \to \pi_G$ in T as follows. By Gödel completeness, it is sufficient to show that any model M of $T + \mathrm{Con}(T)$ satisfies π_G. First, note that π_G is equivalet to $\neg\mathrm{Bew}_T(\overline{\ulcorner \pi_G \urcorner})$, which is also equivalet to $\mathrm{Con}(T + \neg\pi_G)$. Since M satisfies $\mathrm{Con}(T)$, we can make a model M_1 of T over M. So, if M_1 satisfies $\neg\pi_G$, then M shows $\mathrm{Con}(T + \neg\pi_G)$. If M_1 satisfies π_G, M also satisfies π_G since π_G is Π_1 and M is a submodel of M_1.

Some commentaries on Gödel's theorem

- D. Hilbert and P. Bernays, Grundlagen der Mathematik **I-II**, Springer-Verlag, 1934-1939, 1968-1970 (2nd ed.). This gives the first complete proof of the second incompleteness theorem by analyzing the provability predicate.
- P. Lindström, Aspects of Incompleteness, Lecture Notes in Logic 10, Second edition, Assoc. for Symbolic Logic, A K Peters, 2003.
  A technically advanced book, icluding Pour-El and Kripke's theorem (1967) about recursive isomorphisms between recursive theories.
- R.M. Solovay (1976) studied modal propositional logic GL with $\mathrm{Bew}_T(x)$ as modality $\Box$, which is described by
  (1) $\vdash A \Rightarrow \vdash \Box A$,
  (2) $(\Box A \land \Box(A \to B)) \to \Box B$,
  (3) $\Box A \to \Box\Box A$,
  (4) $\Box(\Box A \to A) \to \Box A$

The following are recommended introductory materials.

- T. Franzen, Gödel's Theorem: An Incomplete Guide to Its Use and Abuse (2005).
  On the use and misuse of the incompleteness theorem as a broader understanding of Godel's theorem. A Janpanse translation (with added explanations) by Tanaka (2011).

- P. Smith, Gödel's Without (Too Many) Tears, Second Edition 2022.
  https://www.logicmatters.net/resources/pdfs/GWT2edn.pdf
  Easy to read. The best reference to this lecture.

- K. Tanaka, Math classroom, a graphic guide to the incompleteness theorems (in Japanese), `https://www.asahi.com/ads/math2022/`

Appendix

Jeff Paris

Leo Harrington

- Since Gödel, many researchers were looking for a proposition that has a natural mathematical meaning and is independent from Peano arithmetic, etc.
- Paris and Harrington found the first example in 1977. This is a slight modification of Ramsey's theorem in finite form.
- Following their findings, Kirby and Paris (1982) showed that the propositions on the Goodstein sequence and the Hydra game are independent from PA.
- H. Friedman showed that Kruskal's theorem (1982) and the Robertson-Seimor theorem in graph theory (1987) are independent from a stronger subsystem of second-order arithmetic, and also discovered various independent propositions for set theory.

Thank you for your attention!