The landscape of Gödel's incompleteness theorems

Yong Cheng

School of Philosophy, Wuhan University, China

Celebrating 90 Years of Gödel's Incompleteness Theorems

July 5-9, Nürtingen (Germany), 2021

Outline

- Gödel's incompleteness theorems
- The influence of Gödel's incompleteness theorems
- Different proofs of Gödel's incompleteness theorems
- Generalizations of Gödel's incompleteness theorems
- The intensionality problem of G2
- The limit of the applicability of G1

Gödel's original incompleteness theorems

Gödel proves his first incompleteness theorem in "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I" in 1931 for a certain formal system **P** related to Russell-Whitehead's Principia Mathematica based on the simple theory of types over the natural number series and the Dedekind-Peano axioms.

A theory T is said to be *ω-consistent* if there is no formula $\varphi(x)$ such that $T \vdash \exists x \varphi(x)$, and for any $n \in \omega$, $T \vdash \neg\varphi(\bar{n})$.

Theorem 1 (Gödel's original first incompleteness theorem (G1))

For any formal theory T formulated in the language of **P** *and obtained by adding a primitive recursive set of axioms to the system* **P**, *if T is ω-consistent, then T is incomplete.*

Theorem 2 (Gödel's original second incompleteness theorem (G2))

Let T be a formal theory formulated in the language of **P** *and obtained by adding a primitive recursive set of axioms to the system* **P***, if T is consistent, then the consistency of T is not provable in T.*

In Gödel's 1931 paper, he sketches a proof of the second incompleteness theorem without details and comments in a footnote that it is a corollary of (and in fact a formalized version of) the first incompleteness theorem.

Robinson Arithmetic **Q**

Robinson Arithmetic **Q** is defined in the language $\{\mathbf{0}, \mathbf{S}, +, \cdot\}$ with the following axioms:

Q1 $\forall x \forall y(\mathbf{S}x = \mathbf{S}y \rightarrow x = y)$;

Q2 $\forall x(\mathbf{S}x \neq \mathbf{0})$;

Q3 $\forall x(x \neq \mathbf{0} \rightarrow \exists y(x = \mathbf{S}y))$;

Q4 $\forall x \forall y(x + \mathbf{0} = x)$;

Q5 $\forall x \forall y(x + \mathbf{S}y = \mathbf{S}(x + y))$;

Q6 $\forall x(x \cdot \mathbf{0} = \mathbf{0})$;

Q7 $\forall x \forall y(x \cdot \mathbf{S}y = x \cdot y + x)$.

PA consists of axioms **Q1-Q2**, **Q4-Q7** and the following axiom scheme of induction: $(\phi(\mathbf{0}) \wedge \forall x(\phi(x) \rightarrow \phi(\mathbf{S}x))) \rightarrow \forall x\, \phi(x)$ where ϕ is a formula with at least one free variable x.

The theory **R**

Let **R** be the theory consisting of schemes **Ax**1-**Ax**5 with $L(\mathbf{R}) = \{\mathbf{0}, \overline{n}, +, \cdot, \leq\}$ where $m, n \in \mathbb{N}$.

*Ax*1 $\overline{m} + \overline{n} = \overline{m+n}$;

*Ax*2 $\overline{m} \neq \overline{n}$ if $m \neq n$;

*Ax*3 $\overline{m} \cdot \overline{n} = \overline{m \cdot n}$;

*Ax*4 $\forall x(x \leq \overline{n} \rightarrow x = \overline{0} \vee \cdots \vee x = \overline{n})$;

*Ax*5 $\forall x(x \leq \overline{n} \vee \overline{n} \leq x)$;

A theory T is *locally finitely satisfiable* if every finitely axiomatized subtheory of T has a finite model.

Theorem 3 (Albert Visser)

Suppose T is an r.e. theory with finite signature. Then T is locally finitely satisfiable iff T is interpretable in **R**.

Gödel's incompleteness theorems: modern versions

Theorem 4 (Gödel)

Let T be a recursive enumerable (r.e.) extension of **PA**.

- G1 *If T is ω-consistent, then T is incomplete (there is a sentence θ such that $T \nvdash \theta$ and $T \nvdash \neg\theta$).*
- G2 *If T is consistent, then the consistency of T is not provable in T.*

Theorem 5 (Rosser's first incompleteness theorem)

Let T be a consistent r.e. extension of **R**, *then T is incomplete.*

We will freely use G1 and G2 to refer to both Gödel's first and second incompleteness theorems, and their different versions. The meaning of G1 and G2 will be clear from the context in which we refer to them.

Arithmetization

- The three main ideas in Gödel's proof of G1 are arithmetization, representability and self-reference construction. Let T be a r.e. extension of **PA**.
- Under arithmetization, any formula of the theory can be coded by a natural number (called Gödel's number). We use $\sharp(\phi)$ to denote the Gödel number of ϕ, and use $\ulcorner\phi\urcorner$ to denote the numeral of the Gödel number of ϕ.
- Then we could define some relations on $\mathbb{N}$ which express syntactical properties of T. Define $\mathbf{Prf}_T(m, n)$ iff n is the Gödel's number of a proof of the formula with Gödel number m in T. We can show that $\mathbf{Prf}_T(m, n)$ is r.e..

Representability

- A n-ary relation $R(x_1, \cdots, x_n)$ on $\mathbb{N}^n$ is representable in T iff there is a formula $\phi(x_1, \cdots, x_n)$ such that $T \vdash \phi(\overline{m_1}, \cdots, \overline{m_n})$ if $R(m_1, \cdots, m_n)$ holds; and $T \vdash \neg\phi(\overline{m_1}, \cdots, \overline{m_n})$ if $R(m_1, \cdots, m_n)$ does not hold.
- A key fact we use is: every r.e. relation is representable in **PA**. Let $\mathbf{Proof}_T(x, y)$ be the formula which represents $\mathbf{Prf}_T(m, n)$ in **PA**.
- From the representation formula $\mathbf{Proof}_T(x, y)$, we could define the provability predicate $\mathbf{Prov}_T(x)$ as $\mathbf{Prov}_T(x) \triangleq \exists y \mathbf{Proof}_T(x, y)$.
- $\mathbf{Prov}_T(x)$ satisfies the following conditions:
  (1) If $T \vdash \varphi$, then $T \vdash \mathbf{Prov}_T(\ulcorner \varphi \urcorner)$;
  (2) $T \vdash \mathbf{Prov}_T(\ulcorner \varphi \urcorner) \rightarrow (\mathbf{Prov}_T(\ulcorner \varphi \rightarrow \psi \urcorner) \rightarrow \mathbf{Prov}_T(\ulcorner \psi \urcorner))$;
  (3) $T \vdash \mathbf{Prov}_T(\ulcorner \varphi \urcorner) \rightarrow \mathbf{Prov}_T(\ulcorner \mathbf{Prov}_T(\ulcorner \varphi \urcorner) \urcorner)$.

Self-reference construction

- Gödel constructs a Gödel sentence **G** which asserts its own unprovability in T, i.e. $T \vdash \mathbf{G} \leftrightarrow \neg \mathbf{Prov}_T(\ulcorner \mathbf{G} \urcorner)$.
- Gödel shows that if T is consistent, then $T \nvdash \mathbf{G}$; and if T is ω-consistent, then $T \nvdash \neg \mathbf{G}$.
- Define $\mathbf{Con}(T) \triangleq \neg \mathbf{Prov}_T(\ulcorner 0 \neq 0 \urcorner)$.
- From the above conditions (1)-(3), we can show that $T \vdash \mathbf{Con}(T) \leftrightarrow \mathbf{G}$.
- Thus, G2 holds: if T is consistent, then $T \nvdash \mathbf{Con}(T)$.

Some comments on G1 and G2

- The proof of G1 is constructive: given a consistent r.e. extension T of **Q**, one can effectively find a true Π^0_1 sentence G_T of arithmetic such that G_T is independent of T.
- For Gödel's proof, only assuming that T is consistent does not suffice to show that Gödel sentence is independent of T.
- For any consistent r.e. extension T of **PA**, for each finite sub-theory S of T, $T \vdash \mathbf{Con}(S)$.
- From G2, we cannot get that $\mathbf{Con}(T)$ is independent of T: it is not enough to show that $\neg\mathbf{Con}(T)$ is not provable in T only assuming T is consistent.

Standard provability predicate

Definition 1

- *Unless stated otherwise, we always assume that T is a consistent r.e. extension of* $\mathbf{Q}$.
- *We say that a formula* $\mathbf{Pr}_T(x)$ *is a provability predicate of* T *if* $T \vdash \phi$ *iff* $T \vdash \mathbf{Pr}_T(\ulcorner \phi \urcorner)$.
- *Define the canonical consistency statement* $\mathbf{Con}(T)$ *as* $\neg \mathbf{Pr}_T(\ulcorner 0 \neq 0 \urcorner)$.

Definition 2 (Hilbert-Bernays-Löb derivability condition)

D1 *If* $T \vdash \phi$, *then* $T \vdash \mathbf{Pr}_T(\ulcorner \phi \urcorner)$;

D2 *If* $T \vdash \mathbf{Pr}_T(\ulcorner \phi \rightarrow \varphi \urcorner) \rightarrow (\mathbf{Pr}_T(\ulcorner \phi \urcorner) \rightarrow \mathbf{Pr}_T(\ulcorner \varphi \urcorner))$;

D3 $T \vdash \mathbf{Pr}_T(\ulcorner \phi \urcorner) \rightarrow \mathbf{Pr}_T(\ulcorner \mathbf{Pr}_T(\ulcorner \phi \urcorner) \urcorner)$.

D1-**D3** is called the Hilbert-Bernays-Löb derivability condition.

We say that provability predicate $\mathbf{Pr}_T(x)$ is standard if it satisfies **D2** and **D3**.

Unless stated otherwise, we always assume that $\mathbf{Pr}_T(x)$ *is a standard provability predicate, and* $\mathbf{Con}(T)$ *is the canonical consistency statement defined as* $\neg\mathbf{Pr}_T(\ulcorner \mathbf{0} \neq \mathbf{0} \urcorner)$ *via a standard provability predicate* $\mathbf{Pr}_T(x)$*.*

Lemma 1 (The Diagnolisation Lemma, Gödel)

For any formula $\phi(x)$ *with exactly one free variable, there exists a sentence* θ *such that* $\mathbf{R} \vdash \theta \leftrightarrow \phi(\ulcorner \theta \urcorner)$.

Theorem 6 (Tarski's theorem on the undefinability of truth)

There does not exist a formula $\mathrm{Tr}(x)$ *such that for any formula* ϕ, $\mathbf{R} \vdash \phi \leftrightarrow \mathrm{Tr}(\ulcorner \phi \urcorner)$.

Truth and provability

Definition 3 (**Truth** and **Prov**)

We define **Truth** $= \{\phi \in L(\mathbf{PA}) : \mathfrak{N} \models \phi\}$ *and* **Prov** $= \{\phi \in L(\mathbf{PA}) : \mathbf{PA} \vdash \phi\}$.

In summary, we have:

- **Prov** is definable in the standard model $\mathfrak{N}$ of **PA**, and **Truth** is not definable in $\mathfrak{N}$.
- Both **Prov** and **Truth** are not recursive.
- Both **Prov** and **Truth** are not representable in **PA**.
- **Prov** is arithmetic (a Σ_1 relation), but **Truth** is not arithmetic.

Influence on foundations of mathematics

- Gödel's incompleteness theorems reveal the independence phenomenon which is common in mathematics and logic.
- Gödel's incompleteness theorems show the essential limitation of one given formal system.
- Gödel's incompleteness theorems reveal the essential difference between the notion of "provability in **PA**" and the notion of "truth in the standard model of arithmetic".
- Gödel's incompleteness theorems are a blow to Whitehead-Russell's program for proving that all mathematics (or at least quite a lot of it) could be derived solely from logic in their three-volume Principia Mathematica.
- Gödel's incompleteness theorems have profound influence on the development of Hilbert's program.

Influence on philosophy: Gödel's theorems and the mechanism thesis

- The mechanism thesis claims that human mind can be mechanized.
- A popular misinterpretation of Gödel's theorems is that it implies that human mind cannot be mechanized: the mathematical outputs of the idealized human mind outstrip the mathematical outputs of any Turing machine.
- Lucas and Penrose argue for the anti-mechanism thesis (i.e. human mind cannot be mechanized) based on Gödel's theorem.
- Lucas and Penrose's arguments have been extensively discussed and carefully analyzed in the literature, and are not accepted by most logicians and most philosophers.

Influence on philosophy: Gödel's Disjunction

Gödel did not argue that his theorem implies that human mind cannot be mechanized; instead he argued that his theorem implies a weaker conclusion: Gödel's Disjunction.

The first disjunct (FD) The mind cannot be mechanized;

The second disjunct (SD) There are absolutely undecidable statements in the sense that there are mathematical truths that cannot be proved by the idealized human mind.

Gödel's Disjunction (GD) Either the first disjunct holds or the second disjunct holds.

- For Gödel, he believes that human mind cannot be mechanized, and human mind is sufficiently powerful to capture all mathematical truths; but he thinks that he cannot give a convincing argument for them.
- For Gödel, GD is a mathematically established fact of great philosophical interest which follows from his incompleteness theorem.

Peter Koellner proposes a consistent formal system **DTK**, and shows that:

(1) **DTK** proves GD;
(2) Both Lucas' argument and Penrose's argument for the first disjunct fail in **DTK**;
(3) Both the first disjunct and the second disjunct are independent of **DTK**.

Influence on mathematics

Gödel's proof uses meta-mathematical method and Gödel's sentence has no real mathematical content.

A natural question is then: can we find true sentences not provable in **PA** with real mathematical content?

Harvey Friedman proposed a research program on concrete incompleteness:

the long range impact and significance of ongoing investigations in the foundations of mathematics is going to depend greatly on the extent to which the Incompleteness Phenomena touche normal concrete mathematics.

Concrete incompleteness for **PA**

- Paris-Harrington: Paris-Harrington principle
- Kirby and Paris: The Goodstein sequence, The Hercules-Hydra game
- Kanamori-McAloon: The Kanamori-McAloon principle
- Beklemishev: The Worm principle
- Kirby: The flipping principle
- Mills: The arboreal statement
- Pudlák: P.Pudlák's Principle
- Clote: The kiralic and regal principles
- Weiermann: Variants of Paris-Harrington principle and Goodstein sequence
- Harvey Friedman: Many examples in the book "Boolean relation theory and incompleteness"

Properties of mathematical examples

- Many of these naturally independent principles with real mathematical contents are in fact provably equivalent in **PA** to a certain metamathematical sentence.
- Let $Rfn_{\Sigma_1}(\mathbf{PA})$ denote the sentence which expresses the reflection principle for Σ_1 sentences. People have showed that $\mathbf{PA} \vdash \varphi \leftrightarrow Rfn_{\Sigma_1}(\mathbf{PA})$ for many of the above independent principles φ.
- Many of these independent principles are provable in some fragments of second order arithmetic but are more complex than Gödel's sentence: Gödel's sentence is equivalent to $\mathbf{Con}(\mathbf{PA})$ in **PA**; but many of these principles are not only independent of **PA** but also independent of $\mathbf{PA} + \mathbf{Con}(\mathbf{PA})$.

Concrete incompleteness for Higher-Order Arithmetic

Question 1

Can we find a mathematical theorem expressible in Second-Order Arithmetic but not provable in Second-Order Arithmetic?

Harvey Friedman has found many examples of concrete incompleteness over different fragments of **ZFC** (refer to his book "Boolean relation theory and incompleteness").

Theorem 7 ([1])

There is a concrete mathematical theorem which is expressible in Second-Order Arithmetic, not provable in Second-Order Arithmetic, not provable in Third-Order Arithmetic, but provable in Fourth-Order Arithmetic.

Different proofs of incompleteness

After Gödel, people have found many different proofs of the incompleteness theorems with varied properties:

- Proof via proof theoretic method;
- Proof via recursion theoretic method;
- Proof via model theoretic method;
- Proof via arithmetization;
- Proof via the Diagnolisation Lemma;
- Proof via logical paradox;
- Proof via constructive method;
- Proof only assuming that the base theory is consistent;
- Independent sentences with real mathematical content.

Characteristics of Gödel's proof

Gödel's proof of G1 has the following properties:

- Proof via proof theoretic (meta-mathematical) method;
- Proof via arithmetization;
- Proof without the direct use of the Diagnolisation Lemma;
- Proof "based" on the Liar Paradox;
- Proof via constructive method;
- Proof assuming that the base theory is ω-consistent;
- Gödel's sentence has no real mathematical content.

Incompleteness and logical paradox

- Incompleteness is closely related to paradox.
- Gödel claimed: "Any epistemological antinomy could be used for a similar proof of the existence of undecidable propositions".

  Different proofs of incompleteness theorems via paradox:

Gödel Liar Paradox

Boolos, Chaitin, Kikuchi, Vopenka, Kurahashi, Sakai, Tanaka Berry's paradox

Priest, Cieśliński and Urbaniak Yablo's Paradox

Kritchman-Raz Unexpected Examination Paradox

Cieśliński Grelling-Nelson's Paradox

More in the future

Some definitions

Definition 4

Let T be a consistent extension of $\mathbf{Q}$.

- *T is Σ^0_n-definable if there is a Σ^0_n formula $\phi(x)$ such that n is the Gödel number of some sentence of T if and only if $\mathfrak{N} \models \phi(\overline{n})$.*
- *T is Σ^0_n-sound if for all Σ^0_n sentences ϕ, $T \vdash \phi$ implies $\mathfrak{N} \models \phi$; T is sound if T is Σ^0_n-sound for any $n \in \omega$.*
- *T is Σ^0_n-consistent if for all Σ^0_n formulas ϕ with $\phi = \exists x \theta(x)$ and $\theta \in \Pi^0_{n-1}$, if $T \vdash \neg\theta(\overline{n})$ for all $n \in \omega$, then $T \nvdash \phi$.*
- *T is Π^0_n-decisive if for all Π^0_n sentences ϕ, either $T \vdash \phi$ or $T \vdash \neg\phi$ holds.*

Gödel-Rosser's G1 can be reformulated as follows:

If the theory T is Σ^0_1-definable and consistent, then T is not Π^0_1-decisive.

The following theorems generalize G1 from Σ^0_1-definable theories to Σ^0_{n+1}-definable theories.

Theorem 8 (Kikuchi and Kurahashi)

If T is a Σ^0_{n+1}-definable and Σ^0_n-sound extension of $\mathbf{Q}$, then T is not Π^0_{n+1}-decisive.

Theorem 9 (Salehi and Seraji)

If T is a Σ^0_{n+1}-definable and Σ^0_n-consistent extension of $\mathbf{Q}$, then T is not Π^0_{n+1}-decisive.

The intuitive notion of interpretation

- Interpretability can be accepted as a measure of strength of different theories.
- An interpretation of a theory T in a theory S is a mapping from formulas of T to formulas of S that maps all axioms of T to sentences provable in S.
- $S \trianglelefteq T$ denotes that S is interpretable in T.
- $S \lhd T$ denotes that S is interpretable in T but T is not interpretable in S (i.e. S is weaker than T w.r.t. interpretation).
- S and T are mutually interpretable if $S \trianglelefteq T$ and $T \trianglelefteq S$.

Fact 1 (Folklore)

For theories with finite signature, if S is essentially undecidable and $S \trianglelefteq T$, then T is also essentially undecidable.

If without specification, in the following, we work in theories with finite signature.

The precise notion of interpretation

- Let T be a theory in the language $L(T)$, and S a theory in the language $L(S)$. A translation I of language $L(T)$ into language $L(S)$ is specified by: (1) an $L(S)$-formula $\delta_I(x)$ denoting the domain of I; (2) for each relation symbol R of $L(T)$, an $L(S)$-formula R_I of the same arity; (3) for each function symbol F of $L(T)$ of arity k, an $L(S)$-formula F_I of arity $k+1$.
- A translation I of $L(T)$ into $L(S)$ is an interpretation of T in S if S proves the following facts: (1) for each function symbol F of $L(T)$ of arity k, the formula expressing that F_I is total on δ_I: $\forall x_0, \cdots \forall x_{k-1}(\delta_I(x_0) \wedge \cdots \wedge \delta_I(x_{k-1}) \rightarrow \exists y(\delta_I(y) \wedge F_I(x_0, \cdots, x_{k-1}, y)))$; (2) the I-translations of all axioms of T.
- This describes only *one-dimensional*, *parameter-free*, and *one-piece* translations.

Generalizations of G1 to weak arithmetic

We know that there exists a consistent r.e. weak sub-theory T of **PA** (e.g. Robinson's Arithmetic **Q**) such that each r.e. theory S in which T is interpretable is incomplete.

Let T be a consistent r.e. theory. To generalize G1 to weak arithmetic, we introduce a new notion "G1 holds for T".

Definition 5

G1 *holds for a r.e. theory* T *iff for any consistent r.e. theory* S*, if* T *is interpretable in* S*, then* S *is incomplete.*

Proposition 1

G1 *holds for* T *iff* T *is essentially incomplete iff* T *is essentially undecidable.*

Some key properties of **R**

Theorem 10 (Forklore, many authors)

- *The theory* **R** *is locally finitely satisfiable and hence* **R** *is not finitely axiomatizable.*
- $\mathbf{R} \lhd \mathbf{Q}$ *since* **Q** *is not interpretable in* **R**.
- Σ^0_1*-completeness holds for* **R**.
- *All recursive functions are representable in* **R**, *and hence* **R** *is essentially incomplete.*
- *(Cobham) Any r.e. theory that weakly interprets* **R** *is undecidable.*
- *For each pair* $\langle A, B\rangle$ *of r.e. sets, there exists a formula* $\phi(x)$ *such that* $\mathbf{R} \vdash \phi(n)$ *for* $n \in A \setminus B$, *and* $\mathbf{R} \vdash \neg\phi(n)$ *for* $n \in B \setminus A$.
- *The Lindenbaum algebras of all r.e. theories that interpret* **R** *are recursively isomorphic.*

Summary

In summary, we have the following picture:

- $\mathbf{Q} \lhd I\Sigma_0 + \mathbf{exp} \lhd I\Sigma_1 \lhd I\Sigma_2 \lhd \cdots \lhd I\Sigma_n \lhd \cdots \lhd \mathbf{PA}$, and G1 holds for them.
- The theories $\mathbf{Q}, I\Sigma_0, I\Sigma_0 + \Omega_1, \cdots, I\Sigma_0 + \Omega_n, \cdots, B\Sigma_1, B\Sigma_1 + \Omega_1, \cdots, B\Sigma_1 + \Omega_n, \cdots$ are all mutually interpretable, and G1 holds for them.
- Theories $\mathbf{PA}^-, \mathbf{Q}^+, \mathbf{Q}^-, \mathbf{TC}, \mathbf{AS}, \mathbf{S}^1_2$ and $\mathbf{Q}$ are all mutually interpretable, and G1 holds for them.
- $\mathbf{R} \lhd \mathbf{Q} \lhd \mathbf{EA} \lhd \mathbf{PRA} \lhd \mathbf{PA}$.

For the definition of these weak theories, refer to [2].

Generalizations of G2

Theorem 11 (Löb's theorem)

Let T be a consistent r.e. extension of $\mathbf{Q}$ *and* $\mathbf{Pr}_T(x)$ *be a standard provability predicate. For any sentence* ϕ*, if* $T \vdash \mathbf{Pr}_T(\ulcorner \phi \urcorner) \rightarrow \phi$*, then* $T \vdash \phi$*.*

- Similarly as G1, G2 can also be generalized to some arithmetically definable theories.
- Similarly as G1, G2 can also be generalized to weak arithmetic via interpretation.

Theorem 12 (Pudlák)

There is no consistent r.e. theory S such that $(\mathbf{Q} + \mathbf{Con}(S)) \trianglelefteq S$.

- As a corollary, G2 holds for any consistent r.e. theory that interprets **Q**.
- But it is not true that G2 holds for any consistent r.e. interpreting **R**.

The status of G2

- G1 and G2 are of a rather different nature and scope. Both mathematically and philosophically, G2 is more problematic than G1.
- In the case of G1, we are mainly interested in the fact that *some* sentence is independent of **PA**. We make no claim to the effect that that sentence "really" expresses what we would express by saying "**PA** cannot prove this sentence".
- We can say that G1 is extensional in the sense that we can construct a concrete independent mathematical statement without referring to arithmetization and provability predicate.
- We say that G2 holds for T if the consistency statement of T is not provable in T.
- G2 is essentially different from G1 due to the intensionality problem: whether G2 holds for T depends on how we formulate the consistency statement.

Unless stated otherwise, we assume the following:

- T is a consistent r.e. extension of $\mathbf{Q}$;
- the canonical numbering we use is Gödel's numbering;
- the provability predicate we use is standard;
- the canonical arithmetic formula to express the consistency of the base theory T is $\mathbf{Con}(T) \triangleq \neg \mathbf{Pr}_T(\mathbf{0} \neq \mathbf{0})$;
- the formula representing the set of axioms is Σ^0_1.
- the logic we work on is classic first order logic.

The intensionality of G2

"Whether G2 holds for theory T" depends on the following factors:

(1) the choice of the base theory;

(2) the choice of numberings;

(3) the choice of provability predicate;

(4) the choice of an arithmetic formula to express consistency;

(5) the choice of a specific formula representing (numerating) the axiom set.

(6) the choice of logic we use (to be confirmed)

These factors are not independent, and a choice made at an earlier stage may have effects on the choices available at a later stage.

In the following, unless stated otherwise, when we discuss how G2 depends on one factor, we always assume that other factors are fixed, and only the factor we are discussing is varied.

G2 and the choice of base theory

- An foundational question about G2 is: how much of information about arithmetic is required for the proof of G2. If the base theory does not contain enough information about arithmetic, then G2 may fail.
- Dan Willard has constructed examples of c.e. arithmetical theories that couldn't prove the totality of successor function but could prove their own canonical consistency.
- Fedor Pakhomov defined a weak set theory $H_{<\omega}$ and showed that it proves its own consistency.

G2 and the choice of numberings

- Any injective function γ from a set of $L(\mathbf{PA})$-expressions to $\mathbb{N}$ qualifies as a numbering.
- Gödel's numbering is a special kind of numberings under which the Gödel number of the set of axioms of **PA** is recursive.
- "Whether G2 holds for T" depends on the choice of numberings.
- Grabmayr shows that G2 holds for acceptable numberings; But G2 fails for some non-acceptable numberings.

G2 and the definition of provability predicate

- The consistency statement $\mathbf{Con}(T)$ is usually defined as $\neg\mathbf{Pr}_T(\ulcorner 0 \neq 0 \urcorner)$.
- G2 holds for any standard provability predicate (i.e. satisfying the Hilbert-Bernays-Löb Derivability Condition **D1**-**D3**).
- G2 may also hold for some non-standard provability predicates.
- Define the Rosser provability predicate $\mathbf{Pr}^R_T(x)$ as the formula $\exists y(\mathbf{Prf}_T(x, y) \wedge \forall z \leq y \neg\mathbf{Prf}_T(\dot{\neg}(x), z))$.
- G2 fails for Rosser provability predicate: $T \vdash \mathbf{Con}^R(T) \triangleq \neg\mathbf{Pr}^R_T(\ulcorner 0 \neq 0 \urcorner)$.

G2 and the choice of arithmetic formulas to express consistency

Sergei Artemov argues that in Hilbert's consistency program, the original formulation of consistency "no sequence of formulas is a derivation of a contradiction" is about finite sequences of formulas, not about arithmetization, proof codes, and internalized quantifiers.

Sergei Artemov concludes that G2 does not actually exclude finitary consistency proofs of the original formulation of consistency.

Three arithmetic formulas to express consistency:

(1) $\mathbf{Con}^0(T) \triangleq \forall x(\mathbf{Fml}(x) \wedge \mathbf{Pr}_T(x) \rightarrow \neg\mathbf{Pr}_T(\dot{\neg} x))$;

(2) $\mathbf{Con}(T) \triangleq \neg\mathbf{Pr}_T(\ulcorner \mathbf{0} \neq \mathbf{0} \urcorner)$;

(3) $\mathbf{Con}^1(T) \triangleq \exists x(\mathbf{Fml}(x) \wedge \neg\mathbf{Pr}_T(x))$.

Note that $\mathbf{Con}^0(T)$ implies $\mathbf{Con}(T)$, and $\mathbf{Con}(T)$ implies $\mathbf{Con}^1(T)$.

Definition 6 (The Hilbert-Bernays Derivability Condition)

HB1: *If* $T \vdash \phi \to \varphi$, *then* $T \vdash \mathbf{Pr}_T(\ulcorner \phi \urcorner) \to \mathbf{Pr}_T(\ulcorner \varphi \urcorner)$.

HB2: $T \vdash \mathbf{Pr}_T(\ulcorner \neg\phi(x) \urcorner) \to \mathbf{Pr}_T(\ulcorner \neg\phi(\dot{x}) \urcorner)$.

HB3: $T \vdash f(x) = 0 \to \mathbf{Pr}_T(\ulcorner f(\dot{x}) = 0 \urcorner)$ *for every primitive recursive term* $f(x)$.

- Hilbert-Bernays-Löb derivability condition is sufficient to show that $T \nvdash \mathbf{Con}(T)$ (and hence $T \nvdash \mathbf{Con}^0(T)$), but is not sufficient to show that $T \nvdash \mathbf{Con}^1(T)$.
- Hilbert-Bernays derivability condition is sufficient to show that $T \nvdash \mathbf{Con}^0(T)$, but it is not sufficient to show that $T \nvdash \mathbf{Con}^1(T)$.
- Kurahashi constructed a Rosser provability predicate satisfying the Hilbert-Bernays derivability condition. Thus, the Hilbert-Bernays derivability condition is not sufficient to show that $T \nvdash \mathbf{Con}(T)$.

A general definition of provability predicate

Definition 7

Let $\alpha(x)$ be a formula in $L(T)$.

(1) *Define the formula $\mathbf{Prf}_\alpha(x, y)$ saying "y is the Gödel number of a proof of the formula with Gödel number x from the set of all sentences satisfying $\alpha(x)$".*

(2) *Define the provability predicate $\mathbf{Pr}_\alpha(x)$ of $\alpha(x)$ as $\exists y \mathbf{Prf}_\alpha(x, y)$ and consistency statement $\mathbf{Con}_\alpha(T)$ as $\neg\mathbf{Pr}_\alpha(\ulcorner \mathbf{0} \neq \mathbf{0} \urcorner)$.*

(3) *$\alpha(x)$ is a numeration of T if for any n, $\mathbf{PA} \vdash \alpha(\overline{n})$ iff n is the Gödel number of some $\phi \in T$.*

G2 and the numeration of base theory

Theorem 13 (Gödel)

Let T be any consistent r.e. extension of $\mathbf{Q}$. If $\alpha(x)$ is any Σ_1 numeration of T, then $T \nvdash \mathbf{Con}_\alpha(T)$.

- G2 holds for any Σ_1 numeration of **PA**, but fails for some Π_1 numerations of **PA**.
- Feferman showed that there is a Π_1 numeration $\tau(u)$ of **PA** such that G2 fails: $\mathbf{PA} \vdash \mathbf{Con}_\tau(\mathbf{PA})$.

The limit of the applicability of G1

- Whether a theory is complete depends on the language of the theory. In the languages $L(\mathbf{0}, \mathbf{S})$, $L(\mathbf{0}, \mathbf{S}, <)$ and $L(\mathbf{0}, \mathbf{S}, <, +)$, there are, respectively, recursively axiomatized complete arithmetic theories. For example, Presburger arithmetic is a complete theory of the arithmetic of addition in the language of $L(\mathbf{0}, \mathbf{S}, +)$.
- Recall that G1 holds for some arithmetically definable extensions of $\mathbf{Q}$, but it is not true that any arithmetically definable extension of $\mathbf{Q}$ is incomplete.

The limit of G1 w.r.t. Turing degree

Question 2

A natural question: is there a minimum (or minimal) r.e. theory in some sense for which G1 *holds?*

Define $\overline{\mathrm{D}} = \{S : S <_T \mathbf{R}$ and G1 holds for a r.e. theory $S\}$.

Question 3

(1) *Is* $(\overline{\mathrm{D}}, <_T)$ *well founded? i.e. could we find a minimal theory* S *w.r.t. Turing degree such that* G1 *holds for* S*?*

(2) *Are any two elements of* $(\overline{\mathrm{D}}, <_T)$ *comparable?*

(3) *Is there an infinite descending chain in* $(\overline{\mathrm{D}}, <_T)$*?*

Theorem 14 (Shoenfield)

If $A \subseteq \mathbb{N}$ is a non-recursive r.e. set, then there is a recursively inseparable pair $\langle B, C \rangle$ such that A, B and C have the same Turing degree.

Theorem 15 (Shoenfield)

Let $A \subseteq \mathbb{N}$ be a non-recursive r.e. set. Then there is a consistent axiomatizable theory T having one non-logical symbol which is essentially undecidable and has the same Turing degree as A.

From Theorem 15, for any Turing degree $\mathbf{0} < \mathbf{d} < \mathbf{0}'$, there is a theory $U \in \overline{\mathrm{D}}$ such that U has Turing degree $\mathbf{d}$.

Thus, there is one-to-one correspondence between $\{\mathbf{d} : \mathbf{d}$ is r.e. and $\mathbf{0} < \mathbf{d} < \mathbf{0}'\}$ and $\overline{\mathrm{D}}$.

Hence, the structure $\langle \overline{\mathrm{D}}, \leq_T \rangle$ is as complex as the Turing degree structure of r.e. sets.

Theorem 16

(1) $(\overline{\mathrm{D}}, <_T)$ *is not well founded (or* $\langle \overline{\mathrm{D}}, <_T \rangle$ *has no minimal element).*

(2) $(\overline{\mathrm{D}}, <_T)$ *has infinite many incomparable elements.*

(3) *There is an infinite descending chain in* $(\overline{\mathrm{D}}, <_T)$.

(4) $\langle \overline{\mathrm{D}}, <_T \rangle$ *is dense: for any theory* $A, B \in \overline{D}$ *such that* $A <_T B$, *there is a theory* $C \in \overline{D}$ *such that* $A <_T C <_T B$.

(5) *For any theory* $A \in \overline{D}$, *there exists a theory* $B \in \overline{D}$ *such that* B *is incomparable with* A *under* $\leq_T$.

(6) *Given theories* $A, B \in \overline{D}$ *such that* $A <_T B$, *there is an infinite r.e. sequence of theories* $C_n \in \overline{D}$ *such that* $A <_T C_n <_T B$ *and* C_n *are not comparable.*

(7) *Given theories* $A, B \in \overline{D}$ *such that* $A <_T B$, *any countably partially ordered set can be embedded in* $\langle \overline{\mathrm{D}}, <_T \rangle$ *between* A *and* B.

G1 holds for many theories weaker than **R**

We know that G1 holds for **R** and many weak arithmetics below **PA**. A natural question is:

Question 4

Could we find a theory S such that G1 *holds for S and $S \lhd \mathbf{R}$?*

Definition 8

$\langle S, T\rangle$ is a recursively inseparable pair if S and T are disjoint r.e. sets, and there is no recursive set $X \subseteq \mathbb{N}$ such that $S \subseteq X$ and $X \cap T = \emptyset$.

Theorem 17

For any recursively inseparable pair $\langle A, B\rangle$, there is a theory $U_{\langle A,B\rangle}$ such that G1 *holds for $U_{\langle A,B\rangle}$ and $U_{\langle A,B\rangle} \lhd \mathbf{R}$.*

The limit of G1 w.r.t. interpretation

Define $\mathsf{D} = \{\mathsf{S} : \mathsf{S} \lhd \mathbf{R}$ and G1 holds for a r.e. theory $S\}$.

As a corollary of Theorem 17, D has continuum many elements.

Question 5

(1) *Is* $(\mathsf{D}, \lhd)$ *well founded? i.e. could we find a minimal theory S w.r.t. interpretation such that* G1 *holds for S?*

(2) *Are any two elements of* $(\mathsf{D}, \lhd)$ *comparable?*

(3) *Is there an infinite descending chain in* $(\mathsf{D}, \lhd)$*?*

The interpretation degree structure of r.e. theories extending Robinson's arithmetic **Q** is well known (in fact, a dense distributive lattice under some operations). However, the interpretation degree structure of r.e. theories weaker than Robinson's theory **R** is much more complex and not well known.

We say a r.e. theory U is Turing persistent iff for any consistent r.e. theory V, if $U \subseteq V$, then $U \leq_T V$. As a corollary, we have if U is Turing persistent, then for any r.e. theory V, if $U \unlhd V$, then $U \leq_T V$.

Theorem 18 (Shoenfield, Visser)

For any r.e. set A, there are disjoint r.e. sets B and C with $B, C \leq_T A$ such that for any r.e. D which separates B and C, we have $A \leq_T D$.

Theorem 19

For any r.e. Turing degree $\mathbf{0} < d < \mathbf{0}'$, there exists a Turing persistent theory T_d with Turing degree d such that $T_d \lhd \mathbf{R}$ and G1 *holds for T_d.*

The infimum $A \oplus B$ is defined as follows: $A \oplus B$ is a theory in the disjoint sum of the signatures of A and B plus a fresh 0-ary predicate symbol P. The theory is axiomatised by all $P \to \varphi$, where φ is an axiom of A plus $\neg P \to \psi$, where ψ is an axiom of B.

The supremum $A \otimes B$ is defined as follows: $A \otimes B$ is a theory in the disjoint sum of the signatures of A and B plus two new predicates P_0 and P_1. We have axioms that say that P_0 and P_1 form a partition of the domain and the axioms of A relativised to P_0 and the axioms of B relativised to P_1.

- The degrees of interpretability form a distributive lattice with these two operations.
- For r.e. theories A and B, if A and B both are essentially undecidable, then $A \oplus B$ is also essentially undecidable.

About the structure of $\langle D, \lhd \rangle$, we have the followings which answer some open questions:

- There are countably many elements of D which are incomparable under $\unlhd$.
- There are an descending chain of elements of D under $\lhd$ with countable length.
- If $\langle D, \lhd \rangle$ has a minimal element, then it is also a minimum, and it is not Turing persistent.
- $\langle D, \lhd \rangle$ has no minimal element if we restrict to finitely axiomatized theories.
- $\langle D, \lhd \rangle$ restricted to finitely axiomatized theories is a dense distributive lattice.

The case for infinite signature

We find that whether $\langle \mathrm{D}, \lhd \rangle$ has a minimal element depends on the signature of the language. If the signature of the language is infinite, then $\langle \mathrm{D}, \lhd \rangle$ has a minimum element.

Theorem 20

For any recursively inseparable pair $\langle X, Y \rangle$*, there is a theory* $T_{\langle X,Y \rangle}$ *with infinite signature such that* G1 *holds for* $T_{\langle X,Y \rangle}$ *and* $T_{\langle X,Y \rangle}$ *is interpretable in any first order theory.*

This shows that if without further restrictions, interpretation is not a good notion for comparing essentially incomplete theories since an essentially undecidable theory may be interpretable in a decidable theory.

About the question whether $\langle D, \lhd \rangle$ has a minimal element, it is related to the following factors:

- whether the signature is infinite;
- the complexity of the signature for finite signature;
- the class of theories we consider (e.g. finite axiomatized theory);
- the notion of interpretation we use (e.g. multi-dimensional, piece-wise interpretation with parameters).

Main reference list

- Yong Cheng, Incompleteness for Higher-Order Arithmetic: An Example Based on Harrington's Principle, Springer series: Springerbrief in Mathematics, Springer, 2019.
- Yong Cheng, Finding the limit of incompleteness I. Bulletin of Symbolic Logic, Volume 26, Issue 3-4, December 2020, pp. 268-286.
- Yong Cheng, Current research on Gödel's incompleteness theorem, online in The Bulletin of Symbolic Logic, DOI: 10.1017/bsl.2020.44, 2021.
- Yong Cheng, a manuscript on "Research on Gödel's incompleteness theorems" under work.

Thanks for your attention!