Topics in the Philosophy and Foundations of Mathematics

Lecture 3: Incompleteness

Claudio Ternullo

May 26, 2022

Universitat de Barcelona

Hilbert, 1930

"..Wir müssen wissen, wir werden wissen!"

Hilbert's Programme's Mathematical Goals

Let T be an *infinitary* foundation of mathematics, and F its *finitary* part. Prove (using the resources in F) that:

- T is *consistent*.
- T is *complete*.

To review:

- (Consistency) $T \nvdash \phi \wedge \neg\phi$.
- (Completeness) $(\forall\phi)(T \vdash \phi \vee T \vdash \neg\phi)$.

If T is *consistent*, it doesn't prove a contradiction. If it is *complete* (in the sense shown above!), it is able to prove, for any sentence ϕ, whether it or its negation is a *theorem*.

Completeness

But notice the difference between this version of completeness and the following one:

Completeness

A theory T is *complete* iff:

$$T \models \phi \rightarrow T \vdash \phi$$

Gödel had already proved:

Theorem (Completeness Theorem, [Gödel, 1929])

Given any first-order theory T:

$$T \models \phi \leftrightarrow T \vdash \phi$$

Complete Theories

Completeness is not an unattainable feature of theories. Examples of theories which are complete include:

- The theory F of real-closed fields. $\mathcal{L}_F = \{0, 1, +, \cdot, <\}$
- The theory G of (Abelian) groups. $\mathcal{L}_G = \{0, +\}$.
- The theory DLO of dense linear orders (with no first or last element). $\mathcal{L}_{DLO} = \{<\}$

Formal Arithmetic

We have seen that the theory one may (safely) use to investigate the completeness and consistency of T is some *finitary* part of T, which we denote with F.

In the previous lecture, we characterised F as 'Skolem arithmetic'. A more accurate characterisation of F may be carried out.

For our purposes, our T may just be the (first-order) Peano Axioms (PA). Remember that these were (in a nutshell):

Peano Axioms

1. $N(0)$.
2. $\neg(\exists x)S(x, 0)$.
3. $S(x) = S(y) \to x = y$.
4. (Induction) $F(0) \wedge (\forall x)(F(x) \to F(x+1)) \to (\forall x)F(x)$.

Formal Arithmetic/Cont'd.

$\mathcal{L}_T$ shall consist of:

- A constant symbol 0.
- Three function letters: $f_1(t) = t'$, $f_2(t, s) = t + s$, $f_3(t, s) = t \cdot s$.
- A predicate letter $A(t, s)$ for equality $(=)$.
- Numerals: $\bar{0}, \bar{1}, \bar{2}, \ldots$ (note: $\bar{1} = 0', \bar{2} = 1'$, etc.)

The axioms, as said, are the (first-order) Peano axioms.

(Induction) becomes the induction rule (via MP):

$$\Phi(0), (\forall x)(\Phi(x) \to \Phi(x')) \vdash_T (\forall x)\Phi(x)$$

where 'Φ' is a *schematic* letter standing for any *predicate*.

Enter Gödel: Number-Theoretic Functions

The key reference, here, is [Gödel, 1931].

Gödel considers number-theoretic *functions*, a class of functions which, when taking as arguments natural numbers, output natural numbers.

Let $R(k_1, ..., k_n)$ a number-theoretic relation. This is said to be *expressible* if there exists a formula $\Phi(x_1, ..., x_n)$ such that:

- Whenever $R(k_1, ..., k_n)$ is true, then $\vdash_T \Phi(\bar{k}_1, ..., \bar{k}_n)$;
- Whenever $R(k_1, ..., k_n)$ is false, then $\vdash_T \neg\Phi(\bar{k}_1, ..., \bar{k}_n)$;

Then, he considers number-theoretic functions. These are said to be *representable* if and only if:

- If $f(k_1, ..., k_n) = m$, then $\vdash_T \Phi(\bar{k}_1, ..., \bar{k}_n, \bar{m})$;
- $\vdash_T (\exists! y)\Phi(\bar{k}_1, ..., \bar{k}_n, y)$

Recursivity

Take the initial functions:

- The *zero* function: $Z(x) = 0$.
- The *projection* function: $U_i^n(x_1, ..., x_n) = x_i$ (for all x_i).
- The *successor* function: $S(x) = x + 1$.

and three rules:

- Substitution (Sub).
  $f(x_1, ..., x_n) = g(h(x_1, ..., x_n), ..., h_m(x_1, ..., x_n))$.
- Recursion (Rec).
  - $f(x_1, ..., x_n, 0) = g(x_1, ..., x_n)$.
  - $f(x_1, ..., x_n, y + 1) = h(x_1, ..., x_n, y, f(x_1, ..., x_n, y))$.
- Restricted μ-operator (μ-oper).
  - Suppose $\mu y(g(x_1, ..., x_n, y) = 0)$, the least y such that $g(x_1, ..., x_n, y) = 0$, exists.
  - One can define $f(x_1, ..., x_n) = \mu y(g(x_1, ..., x_n, y) = 0)$.

Recursivity/Cont'd.

- A function is said to be *primitive recursive* (p.r.) if and only if it is obtained from the *initial* functions by a *finite* number of applications of (Sub) and (Rec).
- A function is *recursive* if and only if it is obtained from the *initial* functions by a *finite* number of applications of (Sub), (Rec) and (μ-oper).

Theorem (Gödel)

Assuming T is consistent, the class of all representable (in T) number-theoretic functions is exactly the class of all primitive recursive functions.

Arithmetisation of Syntax

This consists in assigning a natural number to each symbol or string of symbols in $\mathcal{L}_T$.

A Gödel function is a p.r. function $\mathcal{G} : \mathcal{L}_T \to \mathbb{N}$. The value of $\mathcal{G}$ at u is called 'Gödel number of u' (denoted $\ulcorner u \urcorner$, where u is in $\mathcal{L}_T$).

Different symbols will have different Gödel numbers (GN).

Now, consider the following predicates of T:

- $IC(x)$='x is the GN of a constant'
- $FL(x)$='x is the GN of a function letter'
- $PL(x)$='x is the GN of a predicate' letter'

These are all *p.r.* predicates, which shows that T has a *p.r. vocabulary*.

More Primitive Recursive Functions

Now, consider the predicate:

$$PrAx(x)\text{: `}x\text{ is the GN of a proper axiom of } T\text{'.}$$

If $PrAx(x)$ is p.r., then T has a p.r. *axiom set*. It can be proved that T has this property.

As a consequence:

- $Ax(x)$=‘x is the GN of an axiom of T’.
- $Prf(x)$=‘x is the GN of a proof in T’.
- $Pf(x, y)$=‘x is the GN of a proof of the sentence of T whose GN is y’.

are all p.r.

What is F, again?

We have seen that the class of functions *representable* in T is precisely the class of all recursive functions.

One may, then, assume that Hilbert's F is equal to (Primitive) Recursive Arithmetic, the class of number-theoretic functions representable in T.

T, in particular, has a p.r. axiom set (= is *computably axiomatisable*).

So, what Gödel needs to check is whether such a *computably axiomatisable* arithmetical theory T is *complete* and *consistent* by using only F, in turn, reasoning based on p.r. arithmetic.

Fixed Point Lemma

Theorem (Fixed-Point Lemma)

T proves that there exists ψ such that:

$$\psi \leftrightarrow \phi(\ulcorner \psi \urcorner)$$

Proof. Take the p.r. function: $sub(\phi(\ulcorner x \urcorner), m) = \ulcorner \phi(m) \urcorner$. Let $\theta(x) = \phi(sub(x, x))$, and $n = \ulcorner \theta(n) \urcorner$. We put: $\psi = \theta(n)$. Now we have the following equivalences:

$$\psi \leftrightarrow \theta(n)$$

$$\psi \leftrightarrow \phi(sub(n, n))$$

$$\psi \leftrightarrow \phi(sub(\ulcorner \theta(x) \urcorner, n))$$

$$\psi \leftrightarrow \phi(\ulcorner \theta(n) \urcorner)$$

$$\psi \leftrightarrow \phi(\ulcorner \psi \urcorner)$$

Fixed Point Lemma/Cont'd.

As a corollary, one can prove:

Theorem

$T \vdash \psi \leftrightarrow \neg Pr_T(\ulcorner \psi \urcorner)$.

The sentence $\psi \leftrightarrow \neg Pr_T(\ulcorner \psi \urcorner)$ is called *Gödel's sentence* (and is denoted by $\mathfrak{G}$).

$\mathfrak{G}$ says: 'This sentence is such that its Gödel number is not that of a proof in T.'

More simply $\mathfrak{G}$ says: 'This sentence *is not provable*'.

First Incompleteness Theorem

Theorem (First Incompleteness Theorem, [Gödel, 1931])

Assuming the consistency of (the p. r. axiomatisable theory) T, there exists a sentence of T, $\mathfrak{G}$, which isn't provable or disprovable in T.

Proof. Suppose $\mathfrak{G}$ is provable. There is, then, a proof of $\mathfrak{G}$, so we have: $Pr_T(\ulcorner\psi\urcorner)$, the negation of ψ. Contradiction. But T can't prove $\neg\mathfrak{G}$ either, if it is consistent, since this is equal to $Pr_T(\ulcorner\psi\urcorner)$, which, as we have seen, isn't provable (alternatively: $\mathfrak{G}$ is true, since it says that it isn't provable, so there is a *true* sentence of T that isn't provable). $\square$

Second Incompleteness Theorem

Theorem (Second Incompleteness Theorem, [Gödel, 1931])

Assuming the consistency of (the p. r. axiomatisable theory) T, T cannot prove its own consistency (it can't prove $Con(T)$).

Proof. First, formalise 'if T is consistent, then T doesn't prove $\mathfrak{G}$' as a statement of T. So, we have:

$$T \vdash Con(T) \rightarrow \neg Pr_T(\ulcorner \psi \urcorner)$$

but this is equivalent to:

$$T \vdash Con(T) \rightarrow \psi$$

It follows that $T \nvdash Con(T)$ for, otherwise, it would prove ψ (which it doesn't). $\square$

The Rosser Sentence

Rosser formulated a slightly different version of the proof which shows that the notion of *truth* isn't really needed.

Using the Fixed-Point Lemma, he was able to produce the sentence ρ:

'there is a proof of $\neg\rho$ whose Gödel number is less than the Gödel number of a proof of ρ'

Theorem (Rosser, Gödel)

T doesn't prove ρ, and it doesn't prove $\neg\rho$.

Proof. Suppose T proves ρ. Then, there is a proof of $\neg\rho$ with a smaller Gödel number. Suppose T proves $\neg\rho$, which means that 'there is no proof of $\neg\rho$ with a smaller Gödel number'. Now, because of what $\neg\rho$ says, T proves that there exists a proof of ρ (smaller than a proof of $\neg\rho$). In both cases, we derive a contradiction $\square$.

On the Rosser Sentence, Again

It should be noted that Rosser's argument doesn't use:

- The assumption that the theory T is ω-consistent.
- The assumption that T is true.

ω-consistency

A theory T is ω-consistent iff whenever T proves $\neg P(n)$, for each n, it doesn't prove $(\exists x)P(x)$.

This means that:

- The assumption of *truth* is not necessary, so Gödel's argument may just be taken to be purely syntactic.
- The assumption of ω-consistency (in Gödel's original proof) is not necessary, only plain consistency of T is needed.

Final Remarks

One further annotation. Since $\neg Con(T)$ cannot be refuted by T, then it is consistent with T, that is, there is a model M of T such that $M \models \neg Con(T)$.

But notice that $Con(T)$ is a *number-theoretic* statement representable in T, as we know, that is a statement about natural numbers.

If there is a counterexample to it, then there would be natural numbers such that $\neg Con(T)$ and, by the Completeness Theorem, this would be provable. But we have proved that T doesn't prove $\neg Con(T)$.

So, any model of T *must* contain non-standard natural numbers. So, this proves that there are *non-standard* models of arithmetic.

Tarski's Theorem

Suppose we may define a 'truth predicate' $Tr(x)$, which says that 'x is a true sentence of T'.

If such a predicate were p.r., then we could enumerate all the truths of T.

Alfred Tarski, in a celebrated theorem, dashed all such hopes, by showing:

Theorem (Tarski)

There is no p.r. predicate $Tr(x)$ in the language of arithmetic such that $\mathbb{N} \models \psi \leftrightarrow Tr(\ulcorner\psi\urcorner)$, where $Tr(x)$='x is true'.

Proof. It follows from the Fixed Point Lemma that there must be a sentence ψ such that $T \vdash \psi \leftrightarrow \neg Tr(\ulcorner\psi\urcorner)$. $\square$

Remarks on Tarski's theorem

The sentence ψ, this time, says: 'This sentence is such that its GN is not one of a *true* sentence'.

So, ψ just says: 'I am not true'. Since it is provable in T, a contradiction follows immediately.

Tarski's theorem shows that the set of truths of T is not r.e., but the original version of his theorem proves that *truth* isn't definable in any (not necessarily p.r. axiomatisable) theory T.

Tarski's result, thus, posits even stronger limitations on *formal systems*.

Goodstein's Theorem

One could think that Gödel's sentence $\mathfrak{G}$ and Rosser's sentence ρ are just contrived examples of independent statements.

Years later, genuine number-theoretic statements were found, however, which exhibit the incompleteness of PA.

One is related to 'Goodstein sequences'. A Goodstein sequence may be defined as follows:

- Start with a *natural number*, say, N.
- Represent the number as *sum* of *powers* of 2 (extending this procedure to the exponents), to get $N(2)$.
- In $N(2)$, replace all 2s with 3s, and subtract 1 to get $N(3)$.
- In $N(3)$, replace all 3s with 4s and subtract 1, to get $N(4)$.
- ...

The sequence $< N, N(2), N(3), N(4), ... >$ is called *Goodstein sequence*.

Goodstein's Theorem

Theorem (Goodstein)

For any initial N, there is an $n \geq 2$, such that $N(n) = 0$.

Theorem (Kirby, Paris, 1982)

Goodstein's theorem is unprovable in PA.

This can be shown by using 'base-ω'-representations of Goodstein numbers, and these are not available to PA.

Clearly, Goodstein's theorem is provable in stronger theories. For instance, PA+'there exist ϵ_0'.

$\epsilon_0 = \omega^{\omega^{\omega^{\omega^{\cdots}}}}$ is a very important (countable) ordinal, since it represents the *ordinal strength* of PA.

A Digression: Löb's Theorem

Consider Löb's sentence $\mathcal{L}$:

'This sentence is provable'.

Is this sentence true or false? One would naturally be inclined to say that 'it depends'.

Löb famously established that:

Theorem

The Löb sentence is provable, hence it is true. For any sentence ψ we have that:

$$PA \vdash Pr_{PA}(\ulcorner \psi \urcorner) \to \psi, \text{ then } PA \vdash \psi$$

Tower of Incomplete Theories

Incompleteness is a crucial phenomenon of theories which (computably) enumerate their axioms.

If one drops the requirement that the axioms are *p.r.*, just, for instance, *arithmetically definable*, then one may, in fact, obtain theories which prove their own consistency ([Feferman, 1960]).

It should be noted that, e.g., PA+$Con(PA)$ does not prove its own completeness either, so $(Con(PA + Con(PA))$ is independent of the theory.

We may, thus, generate, theories which are 'consistency stronger', such as:

$$PA + Con(PA), PA + Con(PA + Con(PA)), \ldots$$

all of which cannot prove their own consistency, but the consistency of weaker theories.

Two Alternative Programmes

Reverse Mathematics (RM)

RM takes place at the level of second-order number theory. It consists of five big theories: RCA_0, WKL_0, ACA_0, ATR_0, Π^1_1-CA_0 of different strengths. It has been proved that almost all natural theorems of mathematics can be proved (are equivalent to set existence principles) in one of these theories ([Simpson, 2009], [Simpson, 2014]).

Large Cardinals (LC)

LC takes place in the realm of the higher infinite (cardinals provably not existing in ZFC). It has been noticed that most natural (set-theoretic) statements are equiconsistent with ZFC+'there exists a large cardinal κ'. Moreover, LC induces a reduction of incompleteness ([Kanamori, 2009]).

Incompleteness: Moving Beyond

These two programmes seem to express two alternative ways to construe ‘going beyond incompleteness’:

- Although there is no way to fix incompleteness, we may take theories whose consistency is slightly stronger than that of PA as our *foundation*, because, among other things, all of (concrete) maths is expressible in those theories.
- ‘Genuine incompleteness’ depends on our inability to capture the whole of the concept of set; once found the right axioms (*large cardinal axioms* are good candidates), this incompleteness will be strongly reduced. Other forms of incompleteness derive from inherent limitations in ‘finitistic reasoning’.

Lecture's Main Sources

- [Hamkins, 2020], ch. 7
- [Mendelson, 1997], ch. 3
- [Franzen, 2005]

Feferman, S. (1960).
Arithmetization of metamathematics in a general setting.
Fundamenta Mathematicae, 49:35–92.

Franzen, T. (2005).
Gödel's Theorem. An Incomplete Guide to its Use and Abuse.
AK Peters, Natick (MA).

Gödel, K. (1929).
Über die Vollständigkeit der Logikkalkuls.
PhD thesis, University of Vienna.

Gödel, K. (1931).
Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme, i.
Monatshefte für Mathematik und Physik, 38:173–98.

Hamkins, J. D. (2020).
Lectures on the Philosophy of Mathematics.

MIT Press, Cambridge (MA).

Kanamori, A. (2009).
The Higher Infinite.
Springer Verlag, Berlin.

Mendelson, E. (1997).
Introduction to Mathematical Logic.
Chapman and Hall/CRC, Boca Raton (FL).

Simpson, S. (2009).
Subsystems of Second Order Arithmetic.
Cambridge University Press, Cambridge.

Simpson, S. (2014).
Toward objectivity in mathematics.
In *Infinity and Truth*, pages 157–69. World Sci. Publ., Hackensack, NJ.