Week 13: The Limitative Results: Gödel and Tarski

Introduction to the Philosophy of Mathematics
Dr. Neil Barton
neil.barton@uni-konstanz.de

29 January 2020

Recap

Last time we examined *finitism* (of various kinds) *formalism*, and *Hilbert's Programme*.

In particular, Hilbert wanted to do the following:

(1.) Set out arithmetic, analysis and set theory as completely precise formal systems.

(2.) Show by finitary means (i.e. in Primitive Recursive Arithmetic) that no finitary falsehood is derivable in these systems.

If this could be done, he would have shown that particular formal systems for infinitary mathematics do not prove contradictions about natural numbers.

He would thus have validated the use of axiom systems that seem to mention infinite sets in proving facts about natural numbers (even if you don't believe the infinite sets really exist).

Kurt Gödel (1906–1978) was one of the most brilliant logicians of the 20^{th} century.

In 1931, at the age of 25, he would publish his incompleteness theorems that effectively destroyed Hilbert's Programme.

However, as we'll see, Hilbert's Programme lives on in very restricted forms.

1 The Set Up

Let's talk a little bit about the set up underlying Hilbert's Programme. Key were the axioms of $\mathbf{PRA}$. These contain:

(i) Propositional axioms (e.g. $\phi \rightarrow (\psi \rightarrow \psi)$.

(ii) Equality axioms (e.g. for terms τ, σ, $\tau = \tau$; $\sigma = \tau \rightarrow (\phi(\sigma) \rightarrow \phi(\tau))$.

(iii) Rules: Substitution, modus ponens, and free-variable induction with decidable predicates (from $\phi(0)$ and $\phi(y) \rightarrow \phi(s(y))$ conclude $\phi(y)$).

(iv) Principles for defining complex function terms from simple terms in such a way that the value of a function term, for any appropriate input, can be computed (the Primitive Recursive Functions), and symbols for these function terms.

The **Primitive Recursive Functions** (PR) are defined as follows:

(i) The *constant* 0 function is PR.

(ii) The *successor* function $f(n) = n + 1$ is PR.

(iii) The *projection* functions that $p_i^n(x_1, ..., x_i, ..., x_n)$ that always pick out the i^{th} argument place given n many arguments are PR.

(iv) The *composition* of PR functions is also PR. So given PR functions $f, g_1, ..., g_m$ the function:

$$h(x_1, ..., x_n) =$$

$$f(g_1(x_1, ...x_n), ..., g_m(x_1, ..., x_n))$$

is also PR. (*Intuitive idea:* Bundling finitely many PR-functions together is also PR.)

(v) *Primitive recursion* using PR functions is also allowed. Given PR-functions f and g, the function h defined as follows:

- $h(x, 0) = f(x)$
- $h(x, S(y)) = g(x, y, h(x, y))$

is also PR. (*Intuitive idea:* If I can tell you how a function h behaves at 0 in using PR-functions, and given how h behaves at some number n I can tell you using PR-functions how h behaves at $n + 1$, then h as a whole is PR.)

Note: There are *no* quantifiers in **PRA**. However you can have a version with quantifiers, and so I'll use this from now on (it makes stating some things below easier).

PRA represents a theory of functions that are very 'easily' computable in some sense; they can be very explicitly defined.[1]

It's thus often thought that **PRA** is the system that characterises finitism, though there's some debate here.

2 Gödel's Insight and the Incompleteness Theorems

Unfortunately it would take me too long to give a proof of Gödel's Theorems from the ground up (we could have spent most of the semester doing it).

I can give you the key philosophically interesting moving parts though.

Gödel's noticed the following:

[1] There are in fact some computable functions that are not PR. The Ackermann function $A(m, n)$ for instance is a two-place computable function that grows *very* fast (e.g. $A(4, 2) = 2^{65536} - 3$) and has $19,729$ digits). The Wikipedia article on it provides a nice introduction.

Gödel's Insight.

(i) The formal properties Hilbert was interested in (e.g. formulas, proofs, consistency) all concern *strings* of symbols.

(ii) You can code strings of symbols by natural numbers (more below).

(iii) You can then reason about the formal properties *mathematically* by examining what codes exist.

Let's go a little deeper:

What is a *formula* or a *sentence*? It's a particular sequence of symbols.

What are *axioms*? They're particular kinds of sentences, so again, just particular sequences of symbols.

What is a *derivation from some axioms*? It's a particular kind of sequence of symbols where each line is either:

(a) An axiom.

(b) Follows from previous lines by allowed rules of inference.

When is a bunch of axioms *consistent*? When I can't produce a derivation of some contradiction (e.g. $0 = 1$, $(\exists x)x \neq x$, $\phi \wedge \neg\phi$)[2] from my axioms (i.e. there is *no* sequence of a certain kind).

Observation. You can represent any particular string of symbols by a natural number.

[2]Remember all contradictions are equivalent here!

Sketch of the idea. You first assign numbers to your quantifiers, connectives, variables, and terms.[3] So, for instance, we might assign 1 to $\forall$, 2 to $=$, 3 to $($, 4 to $)$, and 11 to x_0 etc. We can then represent a sequence of symbols in a bunch of ways, but one might use *prime decomposition*. So the axiom:

$(\forall x_0)x_0 = x_0$

Would be represented by:

$2^3 \times 3^1 \times 5^{11} \times 7^4 \times 11^{11} \times 13^2 \times 17^{11} \approx 4.65 \times 10^{39}$

as you can see, these codes get big quickly!

We will denote the Gödel code of a syntactic object a by $\ulcorner a \urcorner$ (e.g. the code of a formula ϕ is denoted by $\ulcorner \phi \urcorner$).

Fact. (Gödel) In $\mathbf{PRA}$ you can define predicates of numbers that hold for the syntactic categories of $\mathbf{PRA}$. e.g. there is a predicate $Fmla(x)$ that holds just in case $x = \ulcorner \phi \urcorner$ for some formula ϕ.

Fact. (Gödel) If x is an object in one of our syntactic categories, then $\mathbf{PRA}$ proves that $\ulcorner x \urcorner$ has that property (e.g. for a formula ϕ, $\mathbf{PRA} \vdash Fmla(\ulcorner \phi \urcorner)$).

Philosophical Upshot. $\mathbf{PRA}$ can talk about its own syntax through looking at what numbers are codes of syntactic objects.

In particular we can define a predicate $Prf(x, y)$ that holds just in case y codes a sentence and x codes a proof

[3]A little bit of wizardry is required to avoid double numbering, since you've got infinitely many variables and a symbol for each function term. We'll suppress this here to keep discussion manageable, but check out [Boolos et al., 2007] for the details.

of y from some axioms of $\mathbf{PRA}$.

Fact. (Gödel) If n codes a proof of ϕ in $\mathbf{PRA}$, then $\mathbf{PRA} \vdash Prf(n, \ulcorner\phi\urcorner)$.

Philosophical Upshot. $\mathbf{PRA}$ can compute whether or not a particular number codes a proof of a sentence or not.

Gödel's next step was to show the following *vital* Lemma:

Lemma. (The Diagonal Lemma) For any one-place predicate $\phi(x)$ in the language of $\mathbf{PRA}$, there is a sentence β in the language of $\mathbf{PRA}$ such that:

$\mathbf{PRA} \vdash \beta \leftrightarrow \phi(\ulcorner\beta\urcorner)$

We can now show:

First Incompleteness Theorem. (Gödel) Suppose that $\mathbf{PRA}$ is ω-consistent[4] (this is a slightly stronger assumption than mere consistency). Then there is a sentence G (the *Gödel sentence for* $\mathbf{PRA}$) such that:

$\mathbf{PRA} \vdash G \leftrightarrow \neg(\exists x)Prf(x, \ulcorner G\urcorner)$

and

$\mathbf{PRA} \nvdash G$ and $\mathbf{PRA} \nvdash \neg G$

Why is it the case that $\mathbf{PRA} \nvdash G$?

Rough Idea. Well, if $\mathbf{PRA} \vdash G$, then $\mathbf{PRA} \vdash (\exists x)Prf(x, \ulcorner G\urcorner)$ and so $\mathbf{PRA} \vdash \neg G$, and hence $\mathbf{PRA} \vdash \bot$. If on the other hand $\mathbf{PRA} \vdash \neg G$ then $\mathbf{PRA} \vdash \neg\neg(\exists x)Prf(x, \ulcorner G\urcorner)$, and so $\mathbf{PRA} \vdash G$, and hence $\mathbf{PRA} \vdash \bot$.[5]

So there is at least one sentence G that $\mathbf{PRA}$ cannot prove or refute.

Next we need to define what is is for $\mathbf{PRA}$ to be consistent. That can be done with the following sentence:

$Con(\mathbf{PRA}) =_{df} \neg\exists x Prf(x, \ulcorner 0 = 1\urcorner)$

i.e. There is no code of a proof of $0 = 1$ in $\mathbf{PRA}$.

Fact. $\mathbf{PRA} \vdash Con(\mathbf{PRA}) \rightarrow G$.

Rough idea: If G were derivable in $\mathbf{PRA}$ then $\mathbf{PRA}$ would be inconsistent. So if $Con(\mathbf{PRA})$ holds, then G is not derivable in $\mathbf{PRA}$, but this is exactly what G says.

Second Incompleteness Theorem. (Gödel) If $\mathbf{PRA}$ is ω-consistent then $\mathbf{PRA} \nvdash Con(\mathbf{PRA})$.

Rough idea. If $\mathbf{PRA} \vdash Con(\mathbf{PRA})$, then $\mathbf{PRA} \vdash G$, and so would not be consistent (in the strong way).

3 Philosophical Upshots

Before we have discussion, there's a few points to be raised.

Point 1. As long as your theory $\mathbf{T}$ can be written down as a recursive list of axioms and can represent $\mathbf{PRA}$ (i.e. you can recast all theorems of $\mathbf{PRA}$ as theorems in $\mathbf{T}$) then a version of Gödel's Theorem holds. So other theories also have their own consistency sentences and you can't prove them from within that theory.

This shows that *any* theory able to talk about infinity (even just the natural numbers!) with a decent arith-

[4] A theory $\mathbf{T}$ is ω-inconsistent iff it implies that $\exists n \neg\phi(n)$ but also implies $\phi(n)$ for every standard natural number n.

[5] This really is just a *rough* idea. The assumption of ω-consistency is important.

metic can't prove its own consistency.

Note: Someone asked if there are mathematical systems to which Gödel's Theorem does not apply. Answer: Yes. The important point is that you *need* the codes, and need to be able to treat them mathematically and perform *computations* on them. There are systems of geometry (e.g. a version of Euclidean geometry given by Tarski) for which there is no Gödel-style theorem. *Presburger Arithmetic* (a system which contains only 0, 1, and the binary operation $+$) is decidable and complete. Robinson Arithmetic Q, which only has axioms for successor, addition, and multiplication, and is finitely axiomatisable (no induction!) is incomplete and $Q \nvdash Con(Q)$ (with usual consistency assumptions).

Question. Given all this, was Hilbert wasting his time?

On the one hand yes: His vision could never be realised.

On the other hand, no; his programme gave rise to huge developments that have bearing on mathematics to this day.

What we do now. We know that we can never show consistency. However we can *calibrate* consistency, we can show that certain theories are consistent *relative* to others (i.e. if I assume $Con(\mathbf{T}_1)$ I can prove $Con(\mathbf{T}_2)$).

Question/Optional Exercise. Can anyone think of a theory of arithmetic in the language of $\mathbf{PRA}$ that proves $Con(\mathbf{PRA})$?

4 Questions and Discussion

Optional Exercise 1. Let $Tr_{\mathbf{PA}}$ be a *truth predicate* on the natural numbers such that (for ϕ in the language of $\mathbf{PA}$) $Tr(\ulcorner\phi\urcorner)$ holds iff ϕ is true. Show the following:

Theorem. (Tarski) $Tr_{\mathbf{PA}}$ is not definable in $\mathbf{PA}$.

Hint. Use the Diagonal Lemma and be a Cretian (as per Plato's description).

Optional Exercise 2. Assuming $\mathbf{PRA}$ is ω-consistent is the Gödel sentence G *true*?

Hint. Think about what G says about itself.

Optional Exercise 3. (I'll be mighty impressed if you get this; it's hard and requires some knowledge of models of arithmetic.) Assuming that $\mathbf{PA}$ (or whatever) has the required consistency assumptions, by Gödel's Second we have $\mathbf{PA} \nvdash \neg Con(\mathbf{PA})$, and so $\mathbf{PA} + \neg(Con(\mathbf{PA}))$ is consistent. But $\neg Con(\mathbf{PA})$ says that $\mathbf{PA}$ is inconsistent! So how can $\mathbf{PA} + \neg Con(\mathbf{PA})$ be consistent if it says of itself that its inconsistent?

Hint: There's more than one model of arithmetic.

Question. Does Tarski's Theorem imply that there are no mathematical truths?

Question. Are Gödel's incompleteness theorems a threat to the general idea of a foundation of mathematics?

Question. Do Gödel's Theorems have something fundamentally to do with infinity?

Question. What is meant by the claim "consistency implies existence"?

References

[Boolos et al., 2007] Boolos, G., Burgess, J., P., R., and Jeffrey, C. (2007). *Computability and Logic*. Cambridge University Press.