Week 4. Hilbert's Formalism

Dr. Neil Barton
n.a.barton@ifikk.uio.no
Philosophical Logic and Mathematical Philosophy

10 February 2022

Recap

Last week we talked about *formalism* and *deductivism*.

This week we'll see a very influential programme initiated by David Hilbert that was closely related to formalism.

It was, however, *significantly* impeded by Gödel's Theorems (in my opinion some of the most relevant mathematical theorems for our understanding of the world).

Content warning. Gödel's Incompleteness Theorems are *difficult*.

1 Formalism and Finitism

Hilbert's idea was that the paradoxes should be diagnosed as arising from the notion of infinity.

He therefore proposed that only claims about the finite were acceptable. Call this position **finitism**.[1]

Finitism can be presented in at least three different ways:

Ontological: There is nothing infinite (sets, sequences).

Semantic: Only statements that are both (a) about finite entities, and (b) decidable by finitary means, express real propositions.

Epistemological: Finitary reasoning is secure; non-finitary reasoning introduces uncertainty.

There's lots to talk about (especially in the reading), and I'll leave this for discussion.

Suffice to say, aspects of each appear in Hilbert's thought, and we can think about how these aspects are related to one another.

Part of Hilbert's view concerns *real* vs. *ideal* statements.

Ideal statements have non-finitary content: They imply or presuppose the existence of infinite entities or are not in principle finitarily decidable.

Real statements can be given a decision procedure.

[1]Hilbert scholarship is a little difficult as he seemed to consider more structuralist ideas (the view that mathematics is about structures) in the 1920s. We'll just consider Hilbert the finitist here.

Critical to the *epistemological* part of Hilbert's view is the claim that finitary reasoning is secure, but infinitary reasoning is questionable.

Finitary reasoning contains two key parts:

(1.) Fintary subject matter.

(2.) Finitary methods.

What is finitary subject matter for Hilbert?

Answer. Surveyable arrays of clearly perceptible objects with clearly perceptible shapes and spatial relations.

e.g. Hilbert strokes: ||||||

e.g. A formal language (e.g. symbols, strings of symbols), and their formal properties (e.g. being a formula, being a proof).

These properties are perceptually decidable.

Worry. *Practical limits.* Some arrays are unsurveyable (e.g. a formula with more symbols than atoms in the observable universe).

Solution. Talk schematically. Formulas in the metalanguage can be schemas for formulas in the object language.

e.g. when we say that an instance of $\phi \rightarrow (\psi \rightarrow \phi)$ is an axiom for any ϕ and ψ.

What then are finitary methods?

Answer. There is some debate here, but mostly it is agreed that **Primitive Recursive Arithmetic** (**PRA**) is acceptable.

I won't go through the axioms of **PRA**, but roughly speaking it contains:

(i) 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.

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

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

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

You can think of **PRA** as comprising a fragment of arithmetic where all the functions are easily mechanically computable.

2 Hilbert's Programme

Despite his rejection of the infinite as 'real', Hilbert was nonetheless in many ways positive about our use of talk concerning infinite sets:

> Aus dem Paradies, das Cantor uns geschaffen, soll uns niemand vertreiben können. [Hilbert, 'Über das Unendliche' (1926) p. 170)]

Why is this? Well, Hilbert was an *instrumentalist*, he thought we could use *ideal* statements in proving facts about *real* ones.

For this he initiated what is now known as **Hilbert's Programme**.

The idea was to show by purely finitary means that our ideal mathematical theories (arithmetic, analysis, and set theory) have no false finitary consequences.

His strategy:

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

(2.) Show by finitary means (i.e. in $\mathbf{PRA}$) that no finitary falsehood is derivable in these systems.

The idea was to set out *axiom systems* for infinitary mathematics, and think of our work with ideal mathematics as just computing with these symbols, rather than talking about anything 'real'.

Hilbert hoped to show that these axiom systems, thought of as meaningless symbols (i.e. game formalist-style), would never produce a contradiction.

3 The Incompleteness Theorems

Kurt Gödel (1906–1978) was one of the most brilliant logicians of the 20$^{\text{th}}$ 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 restricted forms.

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 important (and philosophically interesting) moving parts though.

Gödel noticed the following (call this **Gödel's Insight**):

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

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 **PRA** you can define predicates of numbers that hold for the syntactic categories of **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 **PRA** proves that $\ulcorner x \urcorner$ has that property (e.g. for a formula ϕ, $\mathbf{PRA} \vdash Fmla(\ulcorner \phi \urcorner)$).

Philosophical Upshot. **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 of y from some axioms of **PRA**.

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

Philosophical Upshot. **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 **PRA**, there is a sentence β in the language of **PRA** such that:

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

Since it's a bit easier to work with quantifiers (**PRA** doesn't have quantifiers, and tricks are needed to get incompleteness to work) we'll work in Robinson Arithmetic **Q**, which is *very weak* (indeed, it doesn't have any induction and is finitely axiomatisable). It is strong enough to prove the diagonal lemma, however.

We can now show:

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

[2] Remember all contradictions are equivalent here!

[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, Burgess, and Jeffery's *Computability and Logic* for the details.

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

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

and

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

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

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

Note: This can be all done in $\mathbf{PRA}$, but where we need some trickery to mimic the quantifiers (we formalise G via the use of an open formula).

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

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

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

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

Note. Again, $Con(\mathbf{PRA})$ can be formalised in $\mathbf{PRA}$ with suitable trickery to handle the quantifiers. I'll suppress this detail from now on.

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

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).

4 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 arithmetic can't prove its own consistency.

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 e.g. **ZFC** proves $Con(\mathbf{PA})$ (by finding a model for $\mathbf{PA}$).

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

5 Questions and Discussion

5.1 Some optional exercises

Optional Exercise 0. Can anyone think of a theory of arithmetic in the language of $\mathbf{PRA}$ that proves $Con(\mathbf{PRA})$?

Hint: Maybe this theory *extends* $\mathbf{PRA}$ even if it is in the same language.

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. Be a Liar using the Diagonal Lemma.

Optional Exercise 2. Assuming $\mathbf{PRA}$ is ω-consistent is the Gödel sentence G *true* (on the standard model of arithmetic)?

Hint. Think about what G says about itself.

Optional Exercise 3. (Barton's BS)[6] The following argument is bad; explain where it goes wrong. Let $\mathbf{TA}$ be the theory of true arithmetic (i.e. the set of all sentences of arithmetic true on the standard model). $\mathbf{TA}$ is complete. So $\mathbf{TA}$ proves $Con(\mathbf{TA})$.

Hint: Think about what I need to talk about $Con(\mathbf{TA})$.

Optional Exercise 4. (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 $\mathbf{PA}$ (it's not just the standard model) and some models have other 'numbers' in them.

5.2 Discussion

Question. (Fartein/Jens) Hilbert denies the plausibility of actual infinities. How does the incompleteness theorems put forth by Gödel affect actual infinities after Hilbert? Were there made new attempts to "save" actual infinities from Hilbert's objection? Were his objections dismissed? Are Gödel's incompleteness theorems a threat to the general idea of a foundation of mathematics? Did mathematicians simply assume their existence?

Question. (Emma/Jens) Is Hilbert's leap from ideal elements in geometry to ideal numbers in arithmetic valid/legitimate?

Question. (Michel/Jens) Where is the line between what is acceptable and what is unacceptable for the finitist?

Question. Is it acceptable to use ideal theories in proving facts about real statements?

Question. (Magnus/Jens) How is infinity used in physics, and does it matter for philosophy of mathematics? What is the relationship between physical reality and mathematical truth?

[6] So called because I made this mistake.