The Foundations of Mathematics
Lecture Six

Benacerraf's Dilemma

Rob Trueman
rob.trueman@york.ac.uk

University of York

Benacerraf's Dilemma

The Special Properties of Mathematics

- In previous lectures, I have emphasised that mathematics at least *appears* to have some special properties:
  - Mathematical truths are **necessarily** true
  - Mathematical truths can be known **a priori**
  - Mathematical truths can be known with **certainty**
  - Mathematics deals with **infinities**
- The philosophical programmes we have discussed so far were especially concerned with infinity

The Contemporary Perspective

- However, in the contemporary debate, philosophers have largely shifted their attention to the **ontology** of mathematics
  - Do mathematical entities (e.g. numbers, sets) exist?
  - If so, what sort of thing are they, and what sort of relations can we bear to them?
  - If not, then what is mathematics about? How should we understand mathematical claims which *seem* to be about numbers (sets, etc.)?

- Of course, we have also been concerned with these questions all along, but now they take centre stage

- (And although worries about infinity don't disappear either, they move into the background)

Platonism versus Nominalism

- **Platonism** (or *realism in ontology*)
  - Mathematical entities *exist*
  - They are non-physical *abstract objects*: they are outside of time and space, and have absolutely no causal powers
  - They exist *independently* of us

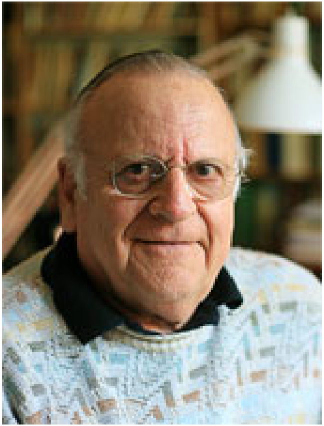
- **Nominalism**
  - Mathematical entities *do not exist*
  - Taken at face value, any theory which appears to describe mathematical objects is *cannot* be true: there are no mathematical entities!
  - The only way that these theories can be true is if there is some way of understanding them as *not really talking about* mathematical objects

Platonism versus Nominalism

- Platonism and nominalism are two ends of a continuum
- It might be possible to find some sort of middle ground
  - Maybe numbers exist, but they aren't abstract objects?
  - Maybe numbers exist, but they aren't independent of us?
- Nonetheless, most contemporary philosophers find themselves leaning more towards one of these extremes or the other
- The debate between platonism and nominalism is primarily driven by **Benacerraf's Dilemma**

Paul Benacerraf

- Benacerraf introduced his dilemma in his 'Mathematical Truth' (reprinted in *B&P*)
- This dilemma became **the** problem to solve in the philosophy of mathematics
- (Benacerraf also wrote another very influential paper, 'What Numbers Could Not Be', which we will come back to in Lecture 9)



Paul Benacerraf

The Dilemma in a Nutshell

- **Semantic Theory**
  - A good philosophy of mathematics should give us a *semantic theory* for mathematics, i.e. a theory of what it takes for a mathematical sentence to be **true**
  - According to Benacerraf, developing a workable semantic theory will push us to **platonism**

- **Epistemology**
  - A good philosophy of mathematics should give us an *epistemology* for mathematics, i.e. a theory of what it takes to **know** a mathematical truth
  - According to Benacerraf, developing a workable epistemology will push us to **nominalism**

Benacerraf's Dilemma

Truths Mathematical and Otherwise

- Consider the following two sentences:
  (1) Some city is bigger than York
  (2) Some prime number is greater than 17
- On the face of it, these two sentences seem to have the same **form**:
  (3) Some F bears relation R to a
- And since they have the same form, it seems sensible to say that they should have the same kind of **truth-conditions**

A Homogenous Semantics

A theory of truth for the language we speak, argue in, theorize in, mathematize in, etc., should [...] provide similar truth conditions for similar sentences. The truth conditions assigned to two sentences containing quantifiers should reflect in relevantly similar ways the contributions made by the quantifiers. Any departure from a theory thus homogeneous would have to be strongly motivated to be worth considering.

(Benacerraf, 'Mathematical Truth' in B&P p. 404)

A Sketch of Tarskian Semantics

- Names **refer** to objects
  - 'a' refers to York
  - 'b' refers to London
- Predicates are **satisfied** by (sequences of) objects
  - x satisfies 'F' iff x is a city
  - $\langle x, y\rangle$ satisfies 'R' iff x is bigger than y
- Quantifiers **generalize** over a domain
  - '$\exists xFx$' is true iff some object in the domain satisfies 'F'
  - '$\forall xFx$' is true iff every object in the domain satisfies 'F'
- **Convention T:** s is true iff P

  (where 's' names a sentence of the object language, and 'P' is a translation of that sentence into the metalanguage)

The Push to Platonism

(1) Some city is bigger than York

- 'York' refers to York
- x satisfies 'city' iff x is a city
- $\langle x, y\rangle$ satisfies 'is bigger than' iff x is bigger than y
- (1) is true iff there is some x s.t. x satisfies 'city' and $\langle x, \text{York}\rangle$ satisfies 'is bigger than'

- On the standard Tarskian semantics, (1) commits us to the **existence** of a city bigger than York

The Push to Platonism

(2) Some prime number is greater than 17

- '17' refers to 17
- x satisfies 'prime number' iff x is a prime number
- $\langle x, y \rangle$ satisfies 'is greater than' iff x is greater than y
- (2) is true iff there is some x s.t. x satisfies 'prime number' and $\langle x, 17 \rangle$ satisfies 'is greater than'

- On the standard Tarskian semantics, (2) commits us to the **existence** of a prime number greater than 17

Combinatorialism

- Benacerraf uses the label **combinatorialism** for any semantic theory that ties the truth-conditions of mathematical sentences to *syntactic* (i.e. combinatorial) features of those sentences
  - **Example:** '$\exists n(Pn \wedge n > 17)$' is true iff $PA \vdash \exists n(Pn \wedge n > 17)$
- We already know that **Gödel's Incompleteness Theorems** will pose a problem for combinatorialism
  - 'Con_{PA}' is true, but $PA \nvdash Con_{PA}$
- However, Benacerraf adds another influential objection…

Face-Value Semantics

- Combinatorialism does not provide a **face-value** semantics for mathematics
  - There are mathematical truths that *appear* to say that certain mathematical entities exist; but according to combinatorialism, they don't really
- Many contemporary philosophers think that there is a presumption in favour of face-value semantics
  - They think that we should only shy away from a face-value semantics for mathematical sentences if there is **linguistic** evidence against it (not mathematical or philosophical evidence)
  - But there is no purely linguistic evidence against the face-value semantics for mathematics

What is Truth?

- Benacerraf presents another kind of objection to combinatorialism:
  - Why should we count being provable from PA as sufficient for being **true**?
- Benacerraf's thought is that Tarski's theory successfully reveals to us what we mean by 'true'
- So what we need is a guarantee that if $\mathrm{PA} \vdash \mathcal{A}$, then $\mathcal{A}$ must be true in Tarski's sense
- But the combinatorialists don't give us that; they just hijack the word 'true' and use it as a label for things you can prove from PA

In Benacerraf's Own Words...

What would make such an assignment of the predicate 'true' the determination of the concept truth? Simply the use of that monosyllable? Tarski has suggested that satisfaction of Convention T is a necessary and sufficient condition on a definition of truth for a particular language. A mere (recursive) distribution of truth values can be parlayed into a truth theory that satisfies Convention T. We can rest with that provided we are prepared to beg what I think is the main question and ignore the concept of translation that occurs in its (Convention T's) formulation...

In Benacerraf's Own Words...

...What would be missing, hard as it is to state, is the theoretical apparatus employed by Tarski in providing truth definitions, i.e., the analysis of truth in terms of the "referential" concepts of naming, predication, satisfaction, and quantification. A definition that does not proceed by the customary recursion clauses for the customary grammatical forms may not be adequate, even if it satisfies Convention T. The explanation must proceed through reference and satisfaction and, furthermore, must be supplemented with an account of reference itself.

(Benacerraf, 'Mathematical Truth' in B&P p. 418)

In Benacerraf's Own Words...

To clarify the point, consider Russell's oft-cited dictum: "The method of 'postulating' what we want has many advantages; they are the same as the advantages of theft over honest toil" (Russell 1919: 71). On the view I am advancing, that's false. For with theft at least you come away with the loot, whereas implicit definition, conventional postulation, and their cousins are incapable of bringing truth. They are not only morally but practically deficient as well.

(Benacerraf, 'Mathematical Truth' in B&P p. 420)

Benacerraf's Dilemma

The Access Problem

- According to platonism, mathematical objects are **abstract** objects
- It is not *entirely* clear what this means, but its generally accepted to entail the following things:
  - Mathematical objects are not located anywhere in space
  - Mathematical objects are not located anywhere in time
  - Mathematical objects have no causal powers
- Placing numbers as such a far remove from ordinary humans leads to the **Access Problem**:
  - How exactly can we know anything about these abstract mathematical objects?

Justified True Belief?

- From Plato until the 1960s, it was generally agreed that knowledge was justified true belief
- Surely we can have justified true beliefs about mathematical entities?
  - It isn't *entirely* clear what 'justified' means, but doesn't maths yields our paradigms of well justified beliefs?
- But in 1963, Gettier presented a number of cases of justified true belief that aren't knowledge

A Gettier Case

- You go to the doctors to be tested for a disease
- A few days later, they bring you back in and give you negative test results

- You come to believe, with good justification, that you do not have the disease
- **BUT!** The doctors mixed up the test results. No one ever noticed, but luckily you actually didn't have the disease
- You have a justified true belief that you do not have the disease, but you still don't *know* that you don't

The Causal Theory of Knowledge

- Benacerraf recommend a version of the **Causal Theory of Knowledge**:
  - To know that P, your belief that P must be caused, in an appropriate way, by the fact that P
  - **Example:** You know that grass is green because you were caused to believe that grass is green by looking at green grass
- The Causal Theory of Knowledge makes it **impossible** to know anything about abstract objects like numbers
  - Abstract objects are non-spatiotemporal
  - But nothing can have causal powers unless it exists in spacetime

A Short-Lived Theory of Knowledge. . .

- The Causal Theory of Knowledge didn't stay popular very long
  - Plenty of philosophers reject it *precisely because* it forbids mathematical knowledge
- But the **Access Problem** can survive the demise of the Causal Theory
- Benacerraf himself makes it clear that the Causal Theory was just one way of making the Access Problem vivid

In Benacerraf's Own Words...

an account of mathematical truth, to be acceptable, must be consistent with the possibility of having mathematical knowledge [...I]n mathematics it must be possible to link up what it is for p to be true with my belief that p. Though this is extremely vague, I think one canse see how [this requirement] tends to rule out [platonism...] For a typical [platonist] account (at least in the case of number theory or set theory) will depict truth conditions in terms of conditions on objects whose nature, as normally conceived, places them beyond the reach of the better understood means of human cognition (e.g., sense perception and the like).

(Benacerraf, 'Mathematical Truth' in B&P p. 409)

The Benacerraf-Field Access Problem

- In *Realism, Mathematics and Modality* (pp.25–30), Hartry Field presents the **Access Problem** without relying on the **Causal Theory of Knowledge**
- **Field's idea:** Forget what mathematicians *know*; their mere *reliability* is already mysterious
  - For the most part, if mathematicians believe mathematical proposition that P, then P
  - We need some explanation of how mathematicians can be so reliable
  - But platonism seems to make any such explanation impossible

In Field's own words . . .

Benacerraf's challenge — or at least, the challenge which his paper suggests to me – is to provide an account of the mechanisms that explain how our beliefs about those remote entities can so well reflect the facts about them. The idea is that if it appears in principle impossible to explain this, then that tends to undermine the belief in mathematical entities, despite whatever reason we might have for believing in them.

(Field, Realism, Mathematics and Modality, pp.25–6)

In Field's own words ...

But special 'reliability relations' between the mathematical realm and the belief states of mathematicians seem altogether too much to swallow. It is rather as if someone claimed that his or her belief states about the daily happenings in a remote village in Nepal were nearly all disquotationally true, despite the absence of any mechanism to explain the correlation between those belief states and the happenings in the village

(Field, Realism, Mathematics and Modality, pp.26–7)

The Push to Nominalism

- To be clear: Benacerraf is not wanting to deny that we know mathematical truths
- He takes it for granted that we **do** have such knowledge
- The problem is that platonism seems to make that knowledge totally mysterious!
- It seems, then, that this gives us a good reason to reject platonism and move towards nominalism
- If we were nominalists, then we would deny that mathematical truths are really about abstract objects, and so would not have to deal with the platonist's Access Problem

Semantics versus Epistemology

- A good **semantics** for mathematics will deal with quantification in mathematical sentences in the same way that it deals with quantification in non-mathematical sentences
- This pushes us towards **platonism**: if there are mathematical truths beginning 'There is...', then mathematical objects exist

- A good **epistemology** for mathematics will make our knowledge of mathematical truths unmysterious
- This pushes us towards **nominalism**: if mathematical truths were about abstract objects, then we would be confronted with the Access Problem

Semantics versus Semantics

- In some ways, it is odd that Benacerraf set up his dilemma as a conflict between semantics and epistemology
- He *could* have set it up as a conflict **within semantics**
- Our ability to **refer** to abstract mathematical objects seems mysterious too!
- We cannot interact with abstract objects, so how can we refer to them?
  - Benacerraf explicitly accepts a **Causal Theory of Reference**
  - Moreover, Causal Theories of *Reference* have proven much longer lived than Causal Theories of *Knowledge*...

Semantics versus Semantics

- Thus we are pulled in two directions in our semantic theorising
- We want our semantics for mathematics to look basically the same as our semantics for everything else...
- ...but we don't want our ability to refer to mathematical objects to be mysterious!

Benacerraf's Dilemma

What Should We Do!?

- **Platonist solutions**
  - Find a way to solve the Access Problem(s)
  - Find a way to excuse not having a solution to the Access Problem(s)

- **Nominalist solutions**
  - Find a way of defending a non-face-value semantics for mathematics
  - Stick with the face-value semantics, even though that makes mathematical claims false by nominalist lights (**fictionalism**)

Gödel's Platonist Epistemology

- Gödel had a two-step epistemology:
  - There are some mathematical truths that we know through quasi-perceptual **intuition**
  - We justify other mathematical truths by their consequences: a new theory is justified if it has theoretical virtues, implies mathematical truths we know by intuition, and implies nothing we know is false by intuition
- Sadly, this isn't really a solution to the epistemological Access Problem: 'intuition' is just a label for the mysterious relation between us and abstracta
- But it *would* be a solution if we said that we *literally* perceive mathematical objects. . .

From Gödel to Maddy

- According to Penelope Maddy in *Realism in Mathematics*, we literally **perceive** sets
- Maddy denies that sets are abstract (in the sense of being non-spatiotemporal)
- Instead, she locates sets where there members are
  - **Example:** The set of everyone in this room is in this room too
- We then justify more complex theories by their (observable) consequences

But *HOW* Do We Perceive Sets?

- It seems that a lot of cognitive processing is involved in seeing ordinary objects
  - We have **object detectors** which let us perceive objects that go over and above the brief glimpses we catch of them
- Maddy's idea is that similar cognitive processing lets us see sets
  - We have **set detectors** which let us perceive sets of objects over and above their members
- We will discuss Maddy's theory in the seminar!

Required Reading for the Seminar

- Benacerraf's 'Mathematical Truth', available through the VLE
- Maddy's *Realism in Mathematics*, chs 1 & 2, available through the VLE