Benacerraf's Mathematical Antinomy

Brice Halimi

Paris West University

Paul Benacerraf, "Mathematical Truth" (*The Journal of Philosophy* 70, 1973)

The main issue is about the right construal of mathematical statements:

- Either they are taken at face value, i.e. as being directly *about* certain objects, according to a *referentialist* semantical account. This is what Benacerraf calls the "semantical" account (= Tarski-Gödel).
- Or they are considered as embedded in a symbolic calculus, and their assertion boils down to their formal derivability from certain axioms. This is what Benacerraf calls the "combinatorial" account (= Hilbert) – insofar as truth values are assigned on the basis of purely syntactic (proof-theoretic) facts.

Benacerraf's example:

(1) There are at least three large cities older than New York.
(2) There are at least three perfect numbers greater than 17.

If one construes (2) as sharing the same kind of truth conditions as (1), then one is committed to a referential parsing that aligns numbers on regular objects. Mathematical truth is just one case of ordinary truth.

But, then, how to explain our epistemic access to such entities that numbers are supposed to be?

On the contrary, if one completely rephrases (2) in formal terms, using the axiomatic system of Peano arithmetic or the classical (second-order) Frege-Russell analysis, say, then one ceases to put (2) on a par with (1), and the epistemic access is no more problematic, since the manipulation of symbols according to rules is obviously available to us.

As Benacerraf puts it: "We need only account for our ability to produce and survey formal proofs." (p. 668)

But, then, how to explain that a mathematical sentence is a truth, beyond being merely a formal theorem?

Two conditions are the standards for any "over-all view" of mathematical truth.

The first condition is

> *[...] the requirement that there be an over-all theory of truth in terms of what it can be certified that the account of mathematical truth is indeed an account of mathematical truth. The account should imply truth conditions for mathematical propositions that are evidently conditions conditions of their truth (and not simply, say, of their theoremhood on some formal system). (p. 666)*

What Benacerraf has in mind is a uniform theory of truth for the language as a whole, including the mathematical language as a mere sub-language.

In Benacerraf's reckoning, only a Tarskian referential semantics can do the job.

Second condition for any "over-all view" of mathematical truth:

> *My second condition on an over-all view presupposes that we have mathematical knowledge. [. . .] The minimal requirement, then, is that a satisfactory account of mathematical truth must be consistent with the possibility that some such truths be knowable. [. . .] An acceptable semantics for mathematics must fit an acceptable epistemology. (p. 667)*

According to Benacerraf, the "combinatorial account" precisely stems from epistemological concerns.

Reducing mathematical knowledge to a formal proof activity avoids having to deal with non-empirical objects.

Benacerraf's dilemma

Benacerraf's point is that you cannot fulfill both conditions at the same time.

- Either you endorse the semantical account, and then you get an account of truth *par excellence*, but no viable epistemology.
- Or you endorse the combinatorial account, and then you have the best epistemology possible, but no account of truth at all.

So you cannot have it both ways: You have to choose between truth and knowledge.

But truth without knowledge or knowledge without truth are both self-defeating options.

The semantical account gives genuine truth conditions, but these are given "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)." (p. 667-668)

In the combinatorial account,

> *[...] there is little mystery about how we can obtain mathematical knowledge. We need only account for our ability to product and survey formal proofs. However, squeezing the balloon at that point apparently makes it bulge on the side of truth: the more nicely we tie up the concept of proof, the more closely we link the definition of proof to combinatorial (rather than semantical) features, the more difficult it is to connect it up with the truth of what is being thus "proved" – or so it would appear. (p. 668)*

Summary of Benacerraf's paper:

- Either a uniform semantics for ordinary language is extended to mathematical language, but then one lapses into platonism;
- Or a reasonable epistemology of mathematical knowledge as a proof activity is put forward, but then no account of mathematical truth other than formal is given.

Benacerraf's text does not provide any clear solution. Hence a dilemma.

Miscellaneous remarks:

- The dichotomy is not one (e.g., what about Frege?).
- The semantical account is very weak (but Benacerraf's paper precisely tilts the scales so as to get an embarrassing balance, and thus a dilemma).
- The "semantical" account is not the only semantical account possible: Benacerraf adds subreptitiously the requirement that the semantics has to be compositional and uniform.
- Explaining the referentiality of a mathematical theory, namely granting that that theory is not simply wheelspinning, does not require to abide by the superficial grammar. (And Benacerraf knows that very well.)
- The epistemology is very naive (very basic empirical logicism).

- The semantical account relies on set theory. But then the actual semantical value of the set-theoretic terms that any Tarskian-style semantics resorts to, should be accounted for. Hence an infinite regress (the problem is only pushed back one square higher up).
- Mathematical theories are neither descriptive nor purely formal. Mathematics evades this duality. Mathematical objectivity is neither "over there" nor fictitious.
- "Benacerraf's dilemma" has been, ENORMOUSLY, the focus of analytic philosophy of mathematics since it has been written.
- It also has put off French philosophers of mathematics as simplistic and out of touch with real mathematics.

Idea: Benacerraf's dilemma evokes Kant's Antinomies of Pure Reason, and more specifically the "mathematical" ones.

Reminder: Both opponents of Kant's two first antinomies (i.e., of the "mathematical" ones) are wrong, whereas both opponents of the two last antinomies (i.e., of the " dynamical" ones) are right (albeit from two different points of view).

Claim: A comparison between Benacerraf's dilemma and Kant's mathematical antinomies is called for by strong analogies, and should be useful, since Kant, in addition to presenting a predicament analogous to Benacerraf's dilemma, does provide a solution for it.

Goal: Harnessing the analogy with Kant so as to transpose Kant's solution to Benacerraf's setting (and thus solve Benacerraf's dilemma).

Mathematical Antinomies of Pure Reason

1. First Antinomy (A426-427/B454-455)
    - Thesis: "The world has a beginning in time, and in space it is also enclosed in boundaries."
    - Antithesis: "The world has no beginning, and no bounds in space, but is infinite with regard to both time and space."
2. Second Antinomy (A434-435/B462-463):
    - Thesis: "Every composite substance in the world consists of simple parts, and nothing exists anywhere except the simple or what is composed of simples."
    - Antithesis: "No composite thing in the world consists of simple parts, and nowhere in it does there exist anything simple."

Comments on the Kantian antinomies

In all Kantian antinomies, each claim is mainly negative: it relies on a reductio ad absurdum and feeds entirely upon the impossibility of the opposite claim.

In the same way, each horn of Benacerraf's dilemma draws its strength only from the predicament of the other.

Each Kantian antinomy is about the concept of something unconditioned with respect to some condition.

There are actually two different ways of conceiving of the unconditioned (A417/B445): either as being the last term of the regressive series of conditions, or as consisting in the whole series itself.

In each antinomy, the thesis epitomizes the first conception, the antithesis the second one.

The thesis seeks a first unconditioned *entity* on which the whole series of conditions depends: It is the reason trying to catch up with the understanding, since the unconditioned is presented as an actual object.

In a reversal from that, the antithesis presents the sum of the conditions in the series as constituting an unconditioned totality: it is the understanding trying to catch up with reason.

As a consequence, cosmological ideas are either too large or too small for the empirical regress sustained by the concepts of the understanding (mismatch between understanding and reason).

[Assume] that the world has no beginning: then it is too big for your concept; for this concept, which consists in a successive regress, can never reach the whole eternity that has elapsed.
Suppose it has a beginning, then once again it is too small for your concept of understanding in the necessary empirical regress. For since the beginning always presupposes a preceding time, it is still not unconditioned, and the law of the empirical use of the understanding obliges you to ask for a still higher temporal condition, and the world is obviously too small for this law.
(A486-487/B514-515)

In each antinomy:

- The thesis aims at some absolute entity (what Kant calls an *Object*, as opposed to a *Gegenstand*).
  Thesis = dogmatism
- The antithesis sticks to the limits of sensible experience: The relationships between appearances and the laws of those relationships are the main focus.
  Antithesis = empiricism

THE ANALOGY (not resemblance) is:

- between
  1. the stress put by empiricism on the understanding (in Kant)
  2. and the stress put by the combinatorial account on mathematical proofs (in Benacerraf)
- between
  1. the stress put by dogmatism on *Objects* (in Kant)
  2. and the stress put by the semantical account on mathematical objects (in Benacerraf).

In other words:

- Kantian understanding (characterized as the faculty of rules) $\rightsquigarrow$ production of formal proofs by mathematics
- Kantian reason (characterized as the quest for unconditioned entities) $\rightsquigarrow$ reference of mathematics to an actual non-empirical denotation.

To get back to the first antinomy:

- Instead of saying that space has a definite extension, Benacerraf's dogmatic (= the semanticist) claims that mathematical terms actually refer to a definite entity.
- Instead of saying that space is boundless, Benacerraf's empiricist (= the combinatorialist) claims that a mathematical object is nothing but the open-ended sum of all the formal proofs that we can produce about it.

Analogy between the Antinomy of Pure Reason and Benacerraf 1973:

Kant	Benacerraf
mathematical antinomy	dilemma
antithetic (each thesis feeds upon the contradiction of the other)	negative argumentation
understanding	"our ability to produce and survey formal proofs" (p. 409)
appearances	symbols
possible experience	admissible inference
series of conditions	deductive chains
law established by the understanding	theorem

Analogy cont'd:

Kant	Benacerraf
reason	truth theory
unconditioned being	direct reference
intellectual intuition	Gödelian intuition
dogmatism (thesis)	realism (semantical conception)
Platonism	platonism (in mathematics)
empiricism (antithesis)	finitism (combinatorial conception)
Epicureanism	Hilbertian formalism

Analogy cont'd:

Kant	Benacerraf
The Idea is either too big or too small for the concept of the understanding (A486/B514).	All analyses "bulge either on the side of knowledge or on the the side of truth" (p. 668).
The antithesis favors knowledge of nature, at the cost of the practical (A469-472/B497-500).	The combinatorial account explains mathematical knowledge, at the cost of mathematical referentiality.
The thesis meets the practical interest of reason, but neglects the investigation of nature.	The semantical account dovetails with the pragmatic need of a unified referential framework, but lets the objects prevail over one's possible access to them.
Solution provided by transcendental idealism	*To be specified*

The detour through Kant's Transcendantal Dialectic calls forth two important points:

- FIRST POINT: "Benacerraf's antinomy" is of the mathematical kind, not of the dynamical one.
- SECOND POINT: Kant's treatment hints at a solution of Benacerraf's dilemma itself, since Kant provided a systematic and extensive solution for the antinomies that can be transposed.

FIRST POINT: Benacerraf's dilemma is analogous to a *mathematical* antinomy

It could seem, indeed, that Benacerraf's dilemma is a dynamical antinomy, as though mathematical truth could be looked at from the point of view of understanding (epistemology) as well as from that of reason (semantics), so that both claims would be legitimate within their respective limits.

Against that view, Benacerraf's dilemma must clearly be understood as an antinomy of the mathematical kind, where both opposite claims are false.

Actually, Benacerraf makes it plain that neither account is philosophically and that both accounts must be overcome.

- Benacerraf, "What numbers could not be" (*Philosophical Review* 74, 1965): The semantical account is proved to be wrong.
- Benacerraf, "Frege: The Last Logicist" (in W. Demopoulos, ed, *Frege's Philosophy of Mathematics*, 1981): The combinatorial account is proved to be wrong.

Benacerraf 1965: about two competing accounts of natural numbers (Cantor's and von Neumann's)

This paper clearly shows that the semantical values of numerical terms like '3' or '17' are not univocal. Even when we are using genuine singular terms in mathematics, their reference is not set unambiguously.

The basic lesson to be learned is that neither account can provide the right semantical value for natural numbers.

In fact, both accounts can be construed as being, precisely, two different interpretations of the same mathematical objects.

Cantor's version and von Neumann's correspond to two different structures C and N for $\mathrm{L} = \{\in, \underline{0}, S\}$:

- C is the L-structure whose domain is $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}, \ldots\}$
- N is the L-structure whose domain is $\{\emptyset, \{\emptyset\}, \{\{\emptyset\}\}, \ldots\}$.

It is a standard set-theoretic result that C and N are *mutually interpretable* (in the model-theoretic sense): Cantor's ordinals are the transitive closures of von Neumann's.

Cantor's version and von Neumann's versions = two different presentations of the same mathematical objects.

A mathematical objects like 3 corresponds in fact to two different items in two different models (e.g., $\{\emptyset, \{\emptyset\}, \{\emptyset, \{\emptyset\}\}\}$ in the model C, as well as $\{\{\{\emptyset\}\}\}$ in the model N), supplemented with the proof that the two models are mutually interpretable and that the interpretations at stake relate the two items to each other.

This is true of "the" natural numbers, which have to be understood as the invariant of a series of equivalent models ("equivalent" in some sense to be specified) – the proof of that equivalence being integral to "the natural numbers" themselves.

Quite generally, a mathematical object emerges as the invariant of the series of all its possible presentations, and intrinsically involves the *proof* that all those presentations are presentations of the same thing. This proof is built into the mathematical object *qua* mathematical.

Conclusion: Objects involve proofs. So the semantical account is inconsistent.

Benacerraf 1981: claims that the whole enterprise of Frege's *Grundlagen* is "first and foremost a mathematical one."

Grundlagen, §3:

> *[. . .] the question [as to whether a proposition is a priori or not] is removed from the sphere of psychology, and assigned, if the truth concerned is a mathematical one, to the sphere of mathematics.*

Benacerraf's comment (p. 55):

> *Since arithmetical propositions are at issue, the question of their justification is properly a matter for mathematics. Therefore, the concepts will be so defined as to make it a properly mathematical question whether some arithmetical judgment is analytic or synthetic, a priori or a posteriori.*

The question as to whether a given mathematical proposition is analytic becomes a mathematical one, not only because "the truth concerned is a mathematical one," but because the way of establishing its analyticity is mathematical.

The recognition of special mathematical truths as being analytical requires to turn proofs themselves into mathematical objects in their own right, which are no less epistemically problematic than the number 3.

Conclusion: Proofs can (have to) be turned into objects. So the combinatorial account is inconsistent as well.

So finally BOTH accounts are inconsistent.

CONCLUSION OF THE FIRST POINT:

Benacerraf's dilemma = *mathematical* antinomy (in the Kantian sense) about mathematics

SECOND POINT: Transposition of Kant's solution to Benacerraf's dilemma

One of the main benefits that can be expected from the analogy made between Kant's *Dialectic* and Benacerraf's dilemma is to provide a clue for a solution of the latter.

Benacerraf's 1965 and 1981 papers already suggest the clue: in Benacerraf's dilemma, both opposite claims are false for the same reason (*mutatis mutandis*) as in Kant:
Both opponents take mathematical items as given in themselves.

- "Space has a bounded extension" (dogmatism) $\rightsquigarrow$ Mathematical objects are what they are, once for all.
- "Space is boundless" (empiricism) $\rightsquigarrow$ A mathematical object is nothing but the illimited sum of all the proofs that we can produce about it and does not exist beyond those proofs.

Kant's solution of the first antinomy:

> *If one regards the two propositions 'The world is infinite in magnitude,' 'The world is finite in magnitude,' as contradictory opposites, then one assumes that the world (the whole series of appearances) is a thing in itself. [...] But if I take away this presupposition, or rather this transcendental illusion, and deny that it is a thing in itself, then the contradictory conflict of the two assertions is transformed into a merely dialectical conflict, and because the world does not exist at all (independently of the regressive series of my representations), it exists neither as an in itself infinite whole nor as an in itself finite whole. It is only in the empirical regress of the series of appearances, and by itself it is not to be met with at all. (A504/B532)*

The whole antinomy of pure reason relies on the false assumption that the objects of experience are given *in themselves*.

Kant's solution transposed: Mathematical objects are not things in themselves.

They are neither full-fledged existing entities nor ideal fictions in the course of a proof.

That mathematical objects are not given in themselves means something quite basic, namely that any mathematical object goes with modes of presentation whose nature depends on the kind of object at stake.

Modes of presentation

Various examples of multiple modes of presentation abund in mathematics, on different scales:

- Natural numbers can be defined as Cantor does, or as von Neumann does.
- The symmetric group $\mathfrak{S}_3$ is often described as the group of permutations on $\{1, 2, 3\}$, but can equally well be defined as the group of permutations on any other three-element set.
- Example in algebra: presentation of a group by generators and relations.

The differences between all those cases do not detract from the occurrence of a same phenomenon, namely the diffraction of a "same" mathematical item into various modes of presentation, without which this mathematical item cannot be grasped, let alone studied.

Such is already, in a way, Benacerraf's diagnosis, as early as 1965, before the dilemma was properly coined:

> *Any purpose we may have in giving an account of the notion of number and of the individual numbers, other than the question-begging one of proving of the right set of sets that it is the set of numbers, will be equally well (or badly) served by any one of the infinitely many accounts satisfying the conditions we set out so tediously. (p. 284)*

A mathematical object can never be given "in itself."

A mathematical structure such as the symmetric group $\mathfrak{S}_3$ never goes without some parametered system which allows one to set ideas (i.e. to kill symmetries) and without whose bias the intended structure cannot be reached, i.e., cognitively handled.

Bringing up the notion of mathematical "mode of presentation" (or "account," to use Benacerraf's term) is a general way to single out the pervasive use of such devices throughout mathematics.

Modes of presentation do not boil down to sub-mathematical conditions of concrete mathematical activity, such as the actual drawing compared to the geometrical theorem. They are truly mathematical in nature.

About the very phrase "mode of presentation"

This term was mentioned by Frege in the context of the distinction that he drew between sense and denotation.

The choice that Frege made of the term "presentation" could certainly be driven back to two opposite sources:

- the work of Franz Brentano in psychology, in particular his book *Psychology from an Empirical Standpoint* (1874), where the notion of "mode of presentation" (*Modus des Verstellens*) is ubiquitous
- the foundation of the theory of group presentations by Walther von Dyck (a student of Felix Klein) in his article "Gruppentheoretische Studien," published in 1882 in the *Mathematische Annalen*.

The Fregean notion of mode of presentation should be construed as both epistemological and logical, both as an epistemic access, and in the mathematical sense of the presentation of a group.

About the notion of mode of presentation

A mathematical object is but the invariant of the open-ended series of all its possible modes of presentation, which themselves are neither purely formal items nor independent semantical units but both (lie in between).

The notion of *mode of presentation* is a tentative notion (imprecise, heterogeneous).
Presentations are certainly many-layered: The mode of presentation of some object can itself be turned into an object w.r.t. some of its own modes of presentation.

One example: $\{\{\{\emptyset\}\}\}$ makes sense only in the context of the axioms of ZFC, but also points to an autonomous item in a model of ZFC.

Another (better) example:
$[a, a^n = 1]$ IS (a presentation of) the cyclic group $\mathbb{Z}/n\mathbb{Z}$
It is neither a mere sequence of symbols, nor a name.

Final transposition of Kant's solution to Benacerraf's dilemma

Just as "[...] the objects of experience are *never* given *in themselves*, but only in experience, and they do not exist at all outside it" (*Critique*, A492/B521), in the same way a mathematical structure never exists outside the series of its presentations, none of which can be privileged as giving what would seem to be "the structure itself."

And just as the unreachable completion of the series of conditions for given appearances is the task prescribed to the understanding as a "regulative principle of reason," in the same way the never-ending exploration of all possible presentations of a "same" mathematical structure is the task, never amenable to completion, that defines mathematics as a discipline.

- The mistake of the combinatorial account is the wrong thesis that mathematical presentations do not present anything.
- The mistake of the semantical account – the mistake of contemporary structuralism (see the "identity problem" raised by Jukka Keränen against Stewart Shapiro) – is the wrong thesis that presentations are mere artefacts as opposed to the mathematical structures "themselves."

A last point

As already mentioned above, Kant explains that, in the first antinomy, the idea of the world is too small for the concept of the understanding in the thesis, and too big for it in the antithesis.

In the case of Benacerraf's dilemma, one could be tempted to say the other way around: that in the thesis (the semantical account) the idea of mathematical objectivity is too big for the understanding, and that in the antithesis (the combinatorial account) it becomes too small.

How to explain this apparent inversion?
In fact, there is none.
The idea of mathematical objectivity, as supplied by the semantical account, is really too small, and the same idea, as supplied by the combinatorial account, really too big.

The yardstick, indeed, is not so much our cognitive power as an open-ended series of presentations that mathematicians come up with in history.

The semantical account closes the process too early (and thus supports too narrow an idea of a mathematical object): The content of a mathematical concept is fixed once for all – although it continues being enriched by new theorems, so that the problem becomes to decide whether one keeps the same object when some really new presentation of it is put forward.

On the contrary, the combinatorial account contends that a mathematical object is already nothing else than the (thereby too big) complete series of all that is and will be proved about it, which now raises the symmetric problem of accounting for genuine ruptures in history of mathematics.

Both opponents neglect the deep historical nature of mathematical objectivity – which should come as no surprise, since the main representants of both sides (Tarski and Hilbert, respectively) foregrounded a mainly logical and anhistorical view of mathematics.

What about the original dilemma?

Not sure that it is solved as Benacerraf intended it to be.

(Not sure that Benacerraf intended it to be solved at all anyway.)

What about the french/analytical divide?

French philosophy of mathematics is naturally more sensitive to the idea of mode of presentation.

This is both the heritage of phenomenology and the result of the study of actual mathematical theories.

French leanings toward the idea of "mathematical architecture" (as opposed to a foundational "linear" picture of mathematics).

Conclusion

1. An analogy can be drawn between Kant's Transcendental Dialectic and Benacerraf's dilemma: dogmatism becomes the semantical account, empiricism becomes the combinatorial account.
   Benacerraf's dilemma can then be reconsidered as a mathematical antinomy about mathematical objectivity.

   The analogy turned out to be remarkably steady and sharp.

2. The main motivation of the analogy remains the prospect of transposing Kant's solution so as to suggest a way of solving Benacerraf's dilemma itself.
   The upshot of the analogy is that a mathematical object can never be given in itself because it consists of the open-ended series of all its possible presentations.

3. The semantical account and the combinatorial one are both wrong because they crystallise an open-ended series into either a closed referent or a complete infinite series.
4. The solution consists in acknowledging that mathematical objects do exist, but proposes to conceive of each mathematical object as a series of equivalent presentations which is always in the making, and which cannot be separated from the proofs explaining that those equivalent presentations that it gathers are indeed equivalent, and how they are.
5. The solution of Benacerraf's dilemma drawn from the analogy with Kant's antinomies calls for a more history-sensitive analysis of mathematical objectivity.