Mathematical Undecidables, Metaphysical Realism, and Equivalent Descriptions

Hartry Field
New York University

1. Metaphysical irrealism in mathematics. The term "metaphysical realism" admits a variety of interpretations; but as applied to mathematics, one natural interpretation is as the doctrine that typical undecidable sentences in established branches of mathematics like set theory have determinate or objective truth value.

The reason for focusing on undecidable sentences is that decidable sentences could easily be held to get determinate truth values "cheaply": their being determinately true or determinately false could be held to just *consist in* (or *result from*) their being provable or refutable in some consistent axiom system each of whose axioms and rules we are disposed to accept. For sentences not so decidable, there is no obvious means by which they could possess determinate truth values "cheaply", i.e. short of a thoroughgoing realist position about the nature of mathematics. Of course, they could *come to have* determinate truth values cheaply: we could conventionally decide to adopt new axioms that settle them. What takes a thoroughgoing realism is to hold that they have determinate truth value now, prior to and independent of any such decision.

Hilary Putnam's attitude towards this sort of metaphysical realism in mathematics has undergone a number of shifts through his long career of important articles on the subject. In Putnam 1967a he critically commented on the view that the continuum hypothesis has a determinate truth value that outruns our ability to decide it, and argued that this view depends on the dubious assumption that there is a determinate totality of all possible subsets of a given infinite set.[1] But in an article published later that year[2] he was ridiculing the opposition to metaphysical realism; after putting into the mouths of the opposition a verificationist argument, he writes

> This "argument" is sometimes taken to show that the notion of a set is unclear. For since the argument "shows" (sic!) that the continuum hypothesis has no truth value and the continuum hypothesis involves the concept of a set, the only plausible explanation of the truth-value failure is some unclarity in the notion of a set. (It would be an interesting exercise to find *all* the faults in this particular bit of reasoning. It is horrible, isn't it?) [Putnam 1967b, p. 52; italics and 'sic's his.]

Later paragraphs make clear that his opposition isn't just to the verificationist argument (which wasn't the argument of 1967a), but to the position of the earlier paper.

By 1980 Putnam seems to have shifted back to something like the 1967a view, and he put forward a more explicit argument for it. (He didn't really endorse this argument, but he did think that reflecting on it undermined "metaphysical realism" in some sense of that phrase, and I *believe* he would have taken that to include the sense above.) The argument is based on the idea that there seems to be nothing in our practice with the notions of 'set' and 'member of' that could single out "the standard model" of set theory as opposed to (typical) "nonstandard" ones.

This way of putting the matter is loose. For one thing, the naive way to understand set theory is to suppose that its quantifiers range over all sets (or all classes, if it admits proper classes); but since there is no set that contains all sets (or class that contains all classes), there isn't such a thing as "the standard model of set theory", at least not on any entirely obvious interpretation. So a better way to put Putnam's point is that there's nothing about our practice that determines that

(i) our restricted quantifier 'all sets' ranges over *all sets*

and

(ii) our predicate '$\in$' applies to a pair <a,b> if and only if a is a member of b.

There are all sorts of possible alternatives to (i) and (ii). Of course, many of these alternatives can be ruled out, as not according with our intentions: this is most obviously so for alternatives that would fail to validate the axioms for sets that we have laid down. But there are plenty of alternatives that validate all of the axioms we have laid down, and in some of these the continuum hypothesis has one truth value and in others another. Putnam seems at first blush to be arguing that there is no way that our practice with the terms 'set' and '$\in$' could determine the semantics of these terms except by restricting the allowable interpretations to those that are consistent with the axioms we've laid down.

Actually, that is only a crude approximation to Putnam's argument. He does *not* argue that consistency with the axioms we've laid down is the only constraint on the interpretations of 'set' and '$\in$'; on the contrary, he implicitly suggests a way in which the empirical applications of mathematics might restrict the range of acceptable interpretations. Although he himself doesn't argue that the empirical applications actually do rule out some interpretations that are consistent with the axioms, I think it is possible to argue persuasively that they do: in particular, they rule out interpretations whose number-theoretic fragments are nonstandard (i.e., in which the quantifier 'only finitely many' comes out nonstandard). If so, all undecidable sentences of number theory (and certain "atypical" undecidable sentences of set theory) do get determinate truth value even assuming Putnam's argument. I've argued this elsewhere (Field 1994, 1998); and a careful reading of Putnam 1980 makes it pretty clear that the point was no news to him. What Putnam argues there (in contrast to the crude approximation in the last sentence of the previous paragraph) is that even making allowance for how the empirical applications of mathematics can constrain the allowable interpretations, there is no way that they could constrain them *enough* to determine truth values for typical undecidable sentences of set theory, such as the continuum hypothesis. His argument that the empirical applications can only do so much seems to me thoroughly convincing, but I don't want to go into these matters again.

For my purposes here, the difference between the cruder argument and the more sophisticated argument won't matter. A key claim of both is that there's no way our practices could determine the extensions of 'set' and/or '∈' sufficiently tightly to give the continuum hypothesis objective truth value.

But as Paul Horwich has noted (Horwich 1998), there is an ambiguity in the notion of determination. In a weak sense of the term, all that is required for our practices to "determine" the extension of a term is that its having that extension supervene on those practices. So one could maintain that reference is determined by our practices (e.g., by our accepting axioms X involving 'set' and '∈') simply by maintaining that anyone who accepted those axioms would inevitably be using the word 'set' for all and only the sets and '∈' for all and only the membership pairs. It is natural to expect, though, that any such claim of supervenience itself needs explanation. (Note that if we are happy with a bare supervenience claim, there is no obvious reason why the axioms X need be true or even approximately true for sets and membership. For all that's been argued, they could be axioms that intuitively seem wholly inappropriate to sets but appropriate instead to kangaroos or harmonic oscillators. Also, for all that's been argued, someone could accept completely analogous axioms that differed only in having different words in place of 'Set' and '∈', and his words would have different extensions: there's no obvious reason why the orthography of our words couldn't be part of the supervenience base.) What Putnam clearly wanted was an illuminating account of *why* using words in the way we use 'set' and '∈' makes them words for all and only the sets and all and only the membership pairs. And there is a very natural stronger sense of 'determines' in which only such an account would be an account of how our practice determines the semantics of these predicates. His point was that while our acceptance of certain axioms involving 'set' and '∈' can (if the axioms are consistent) plausibly be held to determine in the strong sense that the extensions of these predicates are such as to obey these axioms, it isn't in the least clear how accepting those axioms, or anything else about our practice, can (in the strong sense) determine the extension finely enough to settle typical undecidable sentences.

Prima facie, the desire for an account of how semantics is determined in the strong sense is completely reasonable. It is, after all, what we expect for most non-semantic notions (outside of basic physics, which is what we take to be the source of determination). For instance, it's one thing to know that the property of having a gene for hemophilia is weakly determined by (supervenes on) having DNA with a certain feature, and another to know exactly why that feature leads to hemophilia, i.e. to have an account of how, in the strong sense, that feature of the DNA determines the possession of the gene. Lacking such an account of how DNA strongly determines phenotype would be intellectually unsatisfying, and a persuasive argument that there could be no such account would put us in a major intellectual crisis. Why shouldn't we be in a comparable intellectual crisis if we thought that there were objective facts about reference but that they were not strongly determined? If the strong determination of reference can't be maintained, it would seem better to give up the objectivity of reference.

It is natural, then, to assume the following principle:

(P) If no account is possible of how our practice with a term determines (in the strong sense) that it is true of x, or that it isn't true of y, then it can not be objectively (determinately) the case that it is true of x but not of y.

This principle or something like it is assumed in Putnam's argument. I don't say he *believed* the principle: at the very end of the article he renounces the argument he developed in the early sections, and as I understand him, it is Principle (P) that he is renouncing. I'll come back to this. But whatever his ultimate judgement, Principle (P) has considerable initial plausibility, and putting it together with the argument that nothing in our practice can (strongly) determine the extensions of 'set' and '$\in$' finely enough to fix the size of the continuum, we get the conclusion that there can be no objectively correct answer as to what the size of the continuum is.

There is an obvious point in common between Putnam's 1967a argument and his 1980 argument: the 1967a argument was centered on a claimed unclarity in the notion of 'all sets', and in the 1980 argument we get a "semanticized" version of the same thing: there is nothing to determine the semantics of 'all sets' uniquely, so it is indeterminate. Recently Vann McGee (2000) has argued that what makes our quantifiers determinately range over everything is simply their obeying the ordinary rules; to make them range over less than everything, one would need to do something special to restrict them, and the fact that we haven't done that means that our quantifiers determinately range over everything. I'm not fully persuaded, but let's suppose he's right. It still doesn't follow that 'all sets' is determinate, for that is a restricted quantifier whose determinacy is affected by that of 'set'.

Moreover, even if one were to grant that 'all sets' was completely determinate, Putnam's argument would be little affected: the indeterminacy in '$\in$' is all he needs. The intuitive idea behind his argument, after all, is this: suppose, with the mathematical realist, that somewhere outside of space-time there resides a vast array of hunks of platoplasm. Then even if it is somehow determinate which hunks our word 'set' applies to, there is still the question of which pairs of such hunks are in the extension of our symbol '$\in$'. What is there about our use of this term that could settle the matter? The acceptance of axioms of pure set theory could partially constrain it, by ruling out extensions that would make the axioms come out false. The acceptance of impure set theory could add additional constraints, alluded to six paragraphs back, which serve to rule out the *arithmetically* non-standard interpretations. But there don't seem to be any other constraints, and if not, then there is nothing to give determinate truth value to the continuum hypothesis. *Even without questioning the determinacy of 'all sets'*, we have an argument against there being a determinate truth value for typical undecidable sentences of set theory.

Afficionados of second order set theory often argue that Putnam's whole line of reasoning is invalidated by his insistence on using only first order set theory; second order set theory is supposed to avoid the problem, because the standard semantics for it doesn't recognize something as an "interpretation" if the predicate variables don't range over all the subclasses of the domain. But it is a complete illusion to think that you can avoid the problem in this way, as Weston (1976) noted long ago: one might as well "avoid the problem" by using first order set theory but with a stipulation that by 'interpretation' one will mean 'standard interpretation'. I'm not objecting to second order reasoning, any more than to first order set-theoretic reasoning; but it begs all the questions at issue to suppose that second order assertions are all semantically determinate, just as it would beg all the questions at issue to suppose that first order set-theoretic sentences are all semantically determinate.[3]

An interesting recent variant of the use of second order set theory to respond to Putnam's argument is the use of "schematic first order set theory": first order set theory in which the schematic letters (in the replacement and separation schemas) are taken to be extendible as the language evolves. Interesting arguments that this view of schemas is enough to evade Putnam-like problems can be found in McGee 1997, Lavine 1994, Lavine unpublished; and with a bit more qualification, in Shapiro 1991. I have examined these arguments elsewhere (Field 2001, Postscript to Ch 12), and will here simply report my conclusion, which is that they do not get off the ground.[4]

I think, then, that Putnam's argument survives the sort of attacks so far considered: the sort of attacks that do not challenge the basic assumption (P). But is (P) itself a reasonable principle?

2. "Non-realist semantics". Putnam himself challenged (P) at the end of Putnam 1980. The challenge was based on a view he called "non-realist semantics".

I should say at the start that Putnam had put himself in a corner where he had little choice but to challenge (P): he had argued earlier in the paper that (P) leads almost inevitably not merely to such conclusions as that there is no determinate fact as to the size of the continuum, but to somewhat analogous "anti-realist" conclusions about theoretical entities of science, about macroscopic material objects, and about sensations. In my opinion these extensions of the original model-theoretic argument were quite uncompelling (see Lewis 1984, Devitt 1983, Glymour 1982). But I do not want to discuss that matter here; I mention it only to make clear why Putnam had so strong a motivation for challenging (P).

How does he challenge it? (P) could easily be restated as a principle about what makes an interpretation of an expression or mental state "the intended one". What Putnam says in response is that what makes an interpretation in which 'cat' refers to cats intended and one in which 'cat' refers to dogs unintended is simply that we intend 'cat' to refer to cats and not to dogs. Similarly, we intend 'is a member of' to refer to the membership relation; so any argument based on interpretations in which it refers to a relation membership* which need not be the membership relation is irrelevant.

One's first reaction to this idea may well be that it ascribes to us mysterious mental powers. To suppose that we can "intend" cats but not dogs, or membership but not membership*, while also supposing that there is no way of physically grounding that intention, seems to be tantamount to supposing that when one uses the word 'cat' (or the word '$\in$'), Brentanian rays emanate from ones mind and latch onto all and only the cats (or all and only the pairs in which the second is a set and the first is a member of the second), and that these rays ground the intended semantics of the word. But of course Putnam rejects any such view; he calls it an "occult" theory of reference. (And as noted in Lewis 1984, it is in any case hard to see how such a view would really help in escaping the destructive impact that Putnam alleges in his extended model-theoretic argument.)

If Putnam isn't advocating an occult theory of reference, what is he advocating? Obviously the idea is to somehow trivialize the fact that 'cat' in our language refers to cats and 'is a member of' to the membership relation. There is at least one way of doing this: the disquotational theory of reference. The disquotational view trivializes reference for our own language, and reduces reference for other languages to reference for our own language plus translation; it also severely limits the explanatory role of reference, and of related semantic notions such as truth conditions. These things together can be used to explain why weak determination of reference by the physical facts is all we should expect. However, Putnam himself quite explicitly rejects disquotational semantics in many places (e.g. 1983; 1983a, xvii), so it is clear that that is not what he has in mind. Then what *does* he have in mind?

Putnam begins his presentation of his positive view by saying:

> The predicament [i.e., the inability to explain from a naturalistic viewpoint how 'cat' can be true of all and only the cats and how 'is a member of' can be true of all and only the pairs whose second element is a set and whose first element is a member of it] only *is* a predicament because we did two things: first, we gave an account of understanding the language in terms of programs and procedures for *using* the language (what else?); and then secondly, we asked what the possible "models" for the language were, thinking of the models as existing "out there" *independent of any description*. (23-4, italics his)

As the "what else?" indicates, Putnam has no problem with the first of these two steps. Putnam himself had given a persuasive articulation and defense of the procedural account of understanding a couple of years before (Putnam 1978); his judgement there that "the account according to which understanding a language [either a natural language or Mentalese] *consists* in being able to use it (or to translate it into a language one *can* use) is the only account now in the field" (97) is evidently one that he still adhered to.

The step he found problematic, then, was the second, according to which models are "out there independent of any description". But it isn't really the models that are relevant, it is the entities in the domains of the models, and the relations that the models postulate. Is the "problem" with the model-theoretic argument supposed to lie in the assumption that cats and the like exist independent of our descriptions of them?

It is *possible*, if not very charitable, to interpret Putnam as holding that the existence of cats depends on the use of the word 'cat'; and that this dependence gives the referential connection between the word 'cat' and cats. Reference is trivial (for our own words anyway–or for others' words too if their use of language has a part in sustaining the existence of cats and dogs). It is trivial, for instance, that 'cat' applies to all and only the cats and that 'set' applies to all and only the sets. And the reason it is trivial (according to this construal of Putnam) is that cats and sets are shadows of our language-using activity. This constructionist view seems to me to have quite a bit more plausibility for sets than for dogs and cats, and one might consider trying to confine it to that realm. But Putnam is insistent that the model-theoretic argument applies just as much to cats and dogs as to sets, so if we really are to construe him as adhering to constructionism for his way out, we must construe him as a constructionist about cats.

I'm loath to attribute such an implausible doctrine to Putnam (though I must admit that some of his other writings from the period, e.g. Putnam 1981, are hard to interpret otherwise). And much of the end of Putnam 1980 seems thoroughly sensible. For instance:

> To speak as if *this* were my problem, "I know how to use my language, but now, how shall I single out an interpretation?" is to speak nonsense. Either the use *already* fixes the interpretation, or *nothing* can. (p. 24)

This certainly seems right; indeed, it could be expanded to say that *to the extent that* the use doesn't already fix an interpretation, nothing can. Unfortunately, this sensible remark is no solution to the problem: the problem was that it is unclear how the use can completely fix the interpretation, and so, by the principle just enunciated, the interpretation is not completely fixed. That is, we have indeterminacy; and arguably, an unacceptably high degree of indeterminacy. To point out that the argument for indeterminacy leads to indeterminacy is not to give us a means for escaping the argument.

Another thoroughly sensible remark at the end of Putnam 1980, this one more to the point, is that (in accordance with the use theory of understanding) we do have a perfectly good understanding of 'cat', 'set', 'member of', and 'refers'; and

> so we can *say and understand* "'cat' refers to cats". Even though the

> [nonstandard] model referred to satisfies the theory, etc., it is unintended; and we recognize that it is unintended *from the description through which it is given.* (p. 24, italics his).

This seems exactly right to me; and it is just the diagnosis given by the disquotationalist, which leads one to wonder how exactly Putnam's view is different from disquotationalism.

I'm not saying that Putnam is best understood as a disquotationalist: his repeated negative remarks about disquotationalism suggest otherwise. But I suspect that any coherent interpretation of his remarks that avoids ascribing a grossly implausible constructionism, it must be in terms of a position not so far from disquotationalism. At the very least, it must, like disquotationalism, give some kind of explanation of why semantic terms are unlike typical terms of special sciences, in that for semantic terms weak determination by the physical facts is all we need.

In one respect I share with Putnam a worry about disquotationalism. Disquotationalism not only undercuts the sort of argument for indeterminacy considered in the opening section, it appears to undercut the very coherence of indeterminacy; if so, that strikes me as a reductio of the position. I will consider this matter in the next section, and in the following three sections will consider three possible ways for resuscitating the possibility of indeterminacy. (The first attempt at resuscitating it probably presupposes some kind of disquotationalism; the two subsequent ones are compatible with disquotationalism, but very likely are also compatible with whatever alternative way around the argument of Section I Putnam might prefer.) And in Section 7 I will suggest that some of the ideas involved in these defenses of indeterminacy shed some light on one of the most interesting aspects of Putnam's opposition to "metaphysical realism" (not just in mathematics but more generally): his doctrine of equivalent descriptions.

3. Militant disquotationalism. Let's begin with a militant version of disquotationalism, according to which the very idea of indeterminacy is nonsensical. The militant disquotationalist agrees with Putnam's point that we obviously understand terms like 'cat', 'set', 'member of', 'mass' and 'bald' perfectly well, in that we know how to use them; we also understand 'refers', 'true of' and 'true' perfectly well, and understand them to be governed by the schemas

‘t’ refers to t and nothing else, if t exists, and otherwise refers to nothing

‘F’ is true of all and only those things that are F (for 1-place predicates); ‘R’ is true of all and only those pairs <x,y> such that xRy (for 2-place predicates); etc.

and ‘p’ is true if and only if p.

She then claims both (A) that this removes the motivation for positing indeterminacy, and (B) that it shows the idea of indeterminacy to be incoherent.

The point is not merely that disquotationalism makes it incoherent to give up these schemas, and undermines the motivations for doing so: any sophisticated advocate of indeterminacy can agree that the schemas are sacrosanct, i.e. that their instances must be regarded as determinately true. How? By supposing that words like ‘refers’, ‘true of’ and ‘true’ are themselves indeterminate, and their indeterminacy is “correlative to” that of each indeterminate object level term: there are legitimate interpretations in which the extension of ‘bald’ is B_1 and in which <‘bald’,x> is in the extension of ‘true of’ iff x is in B_1, and other legitimate interpretations in which the corresponding things hold for a set B_2 distinct from B_1, but none for which the extension of ‘bald’ is B_1 and in which <‘bald’,x> is in the extension of ‘true of’ iff x is in B_2 (or the analog with B_1 and B_2 reversed). So the determinate truth of (the instances of) the schema

‘F’ is true of all and only the Fs

doesn’t imply the determinacy of whether the instances of ‘F’ are true, and doesn’t imply the schema

‘F’ is determinately true of all and only the Fs.

Similarly, the determinate truth of the schema

‘p’ is true if and only if p

doesn’t imply the schema

‘p’ is determinately true if and only if p.

If the issue then isn't about the schemas, what is it? Let's begin with point (A). The motivation for indeterminacy was that it is hard to see how any theory of reference for 'Set' and '$\in$' (in terms of which '$2^{\aleph_0}$' and '$\aleph_1$' are defined) could determine the reference sufficiently to make one rather than the other of '$2^{\aleph_0} = \aleph_1$' and '$\neg(2^{\aleph_0} = \aleph_1)$' determinately true (given e.g. that no combinations of the axioms we are disposed to accept decide the matter). Similarly, it's hard to see how any theory of reference for 'rich' or 'bald' could determine of each person, no matter how apparently borderline, whether he's rich and whether he's bald. But a disquotationalist insists that we don't need theories of reference in the sense required, i.e., theories of *in virtue of what* names refer to what they refer to, or *in virtue of what* atomic predicates are true of what they are true of. But then it's hard to see how lack of such theories–or even, arguments for the impossibility of such theories–should give us reason to believe in indeterminacy.

I think point (A) is largely correct: more specifically, I agree that disquotationalism undermines any motivation for indeterminacy *based on the theory of reference*. The basic idea of the disquotationalist doctrine is that theories of reference don't serve the philosophical purpose that they have been assumed to serve, and in particular are not needed to ground claims of determinateness of reference. So I'm inclined to agree with the disquotationalist that arguments for indeterminacy based on the theory of reference need to be either recast or abandoned. (The issue of how exactly to recast them, or what arguments might be used in their place, is a pressing one, but beyond the scope of this paper.)

But the more dramatic claim of militant disquotationalism–the part that makes it militant–is (B). The argument for (B) is as follows: Consider any purported example of indeterminacy, e.g. of truth value: say the continuum hypothesis, $2^{\aleph_0} = \aleph_1$. The truth schema implies both that '$2^{\aleph_0} = \aleph_1$' is true iff $2^{\aleph_0} = \aleph_1$ and that'$\neg(2^{\aleph_0} = \aleph_1)$' is true iff $\neg(2^{\aleph_0} = \aleph_1)$. By classical logic it follows that either '$2^{\aleph_0} = \aleph_1$' is true or '$\neg(2^{\aleph_0} = \aleph_1)$' is true (and not both). But this seems to eliminate indeterminacy in any interesting sense. Admittedly, we can stipulate *senses* of indeterminacy that aren't ruled out: e.g., unknowability. But the *intended* sense of indeterminacy of truth value is something like: there is no fact of the matter. And it would seem that if '$2^{\aleph_0} = \aleph_1$' is true (and '$\neg(2^{\aleph_0} =$

$\aleph_1$)' isn't), then there is a fact of the matter as to which is true, viz. that it's '$2^{\aleph_0} = \aleph_1$'; similarly, if '$\neg(2^{\aleph_0} = \aleph_1)$' is true (and '$2^{\aleph_0} = \aleph_1$' isn't), then there is a fact of the matter as to which is true, viz. that it's '$\neg(2^{\aleph_0} = \aleph_1)$'. So either way, there's a fact of the matter. By similar reasoning, there's a fact of the matter as to whether Joe is rich, or bald, or whatever, no matter what the state of his finances or his scalp may be.

It is, I think, extraordinarily difficult to accept the conclusion of this argument. It is, nonetheless, not at all easy to find a convincing reply.

4. Leeds' way out. I will now consider three ways to modify militant disquotationalism to make it more believable.

The first is due to Stephen Leeds (Leeds 1997). Begin by noticing that the schemas for reference, truth etc. apply in the first instance to our own language; so if the determinacy of reference and truth is to be grounded in these schemas, then all we directly get is that it is determinate what *our own terms* refer to and whether *our own sentences* are true. The usual disquotationalist treatment of foreign terms and sentences is via translation; so as long as translation can be indeterminate, we are at least able to recognize an indeterminacy in foreign terms and sentences. And Leeds does so: he says, for instance, that it is indeterminate how to translate certain terms from past theories into our own language, and so it's indeterminate what they refer to. This is the first of his two steps to weakening the militant disquotationalist view. The second step is to suppose that we can make a context-relative distinction between *serious* language and *non-serious* language.[5] In our serious moods, we accept the disquotation schema only for serious language; if it is indeterminate how to translate non-serious language into serious, then reference is indeterminate even for non-serious parts of our own language.

A worry that might be raised about Step 1 of this proposal is whether Leeds can allow the appeal to indeterminacy of translation. 'Synonymous', after all, is a term of our language; but Leeds apparently accepts the disquotationalist view that positing indeterminacy *in our own language* makes no sense, so it follows that we must regard our term 'synonymous' as perfectly determinate. So how can there be indeterminacy of translation?

Although Leeds doesn't consider this objection, I think it's clear what his reply ought to be: that we should insist that the concept of synonymy (in its interlinguistic applications) is not to be included in serious language. By doing so, we avoid serious commitment to the claim that either the foreign term w is synonymous with our term w* or it isn't, so we avoid the conclusion that it is determinately synonymous or determinately non-synonymous with a given term of ours. If the concept of interlinguistic synonymy is so disavowed, what is translation? In translating a term we aren't making any claim about its synonymy to one of our terms, we are simply correlating it with one of our terms for various purposes. Indeterminacy of translation simply means that there is no best policy for translating w; or better, it simply means that given any policy that translates it as w*, there is another policy, neither better nor worse, that gives it an incompatible translation w**. Indeterminacy of translation so understood seems to raise no problems for the disquotationalist.[6]

Another objection to Step 1 is that we don't need to translate foreign terms into our language, we can incorporate them. Suppose we're trying to understand an earlier theory involving a term that seems to have no best translation into modern language: say 'Kraft' as used in 18th century German physics (Leeds' own example; it is sometimes translated as 'energy' and sometimes as 'force'). Rather than translating it, we may simply adopt it into our language for purposes of evaluating or commenting on the theory. But then when we do so, we have lost any means for regarding the term as indeterminate. But this objection too is handled by Step 2 of the proposal: although we incorporate 'Kraft' into our language, we don't take it fully seriously, so there is no problem in saying that it is indeterminate what it stands for.

Not only does Step 2 handle the objections above, it also provides a way for recognizing ordinary vagueness in our own language. It's true that insofar as we regard 'bald' as part of serious language we must say that there are no borderline cases: everyone is either determinately bald or determinately not bald, in any but an irrelevant sense of 'determinately', for in the only sense that is relevant, the term is redundant as applied to our serious language. But we don't in all contexts regard 'bald' as part of our serious language, and insofar as we don't we can recognize indeterminacy as resulting from an indeterminacy as to how to translate it into serious terms.

I'm fairly sympathetic to Leeds' view of indeterminacy (which is not to say that I'm sure it is ultimately defensible). Supposing for the sake of argument that it's right, what consequences does it have for indeterminacy about sentences like the continuum hypothesis? In my view, it allows for their indeterminacy. Of course, insofar as we regard the standard language of set theory as part of our serious language, we can't make sense of the idea that the continuum hypothesis is indeterminate, any more than we can regard Joe as a borderline case of baldness insofar as we regard 'bald' as in our serious language. But there are plenty of alternatives to regarding the language of set theory as part of serious language. For instance, we can take the language in some sort of fictionalist spirit. And even if we are platonists in the sense of believing (in the fullest and most serious sense) that there are infinitely many non-physical eternally existing objects and that set theory is best understood as having them (or some of them) as its subject matter, we needn't understand the language of set theory as meeting the most demanding standards of seriousness. Maybe when being REALLY serious, we just say, without set-theoretic vocabulary, (i) that there are infinitely many non-physical eternally existing objects (which can be done without the term 'infinitely'), and (ii) that standard set theory is consistent (in, say, the modal sense of consistency discussed in Field 1991, which doesn't presuppose set theory). We then translate the language of set theory into the theory consisting of (i) and (ii), by taking 'set' to be true of some or all of the non-physical eternally existing objects and by interpreting 'member of' in any way that makes the usual axioms come out true. There are multiple ways of doing this, and different ones make different sentences about the size of the continuum come out true. So we get a multiplicity of translations between the language of set theory and the fundamental platonistic theory, and the continuum hypothesis comes out lacking in determinate truth value.[7] I think then that even Leeds' close relative of militant disquotationalism leaves room for the idea that we shouldn't be "metaphysically realist" about mathematical undecidables.

5. Abandoning excluded middle. Leeds' view is only one of several responses to the militant disquotationalist argument against indeterminacy. Another more obvious one is to reject the supposition in the argument that the law of excluded middle holds for indeterminate language. Putnam once (1983) recommended the use of intuitionist logic for handling one form of indeterminacy, viz. vagueness. Much more promising, I think, is the use of a logic based on the strong Kleene 3-valued truth tables. Even better is the use of a logic based on arbitrary complete deMorgan lattices (complete distributive lattices that validate the deMorgan laws and the principle that $\neg\neg A$ is equivalent to A). The points in a deMorgan lattice are naturally thought of as "degrees of truth". Different such lattices represent different assumptions about how many such degrees of truth there are, and how they are ordered: e.g., whether the order is linear. (The Kleene semantics is the special case that involves only one degree of truth intermediate between 0 and 1.)[8]

It's clear that with excluded middle rejected, Argument (B) for the incoherence of indeterminacy (given at the end of Section 3) collapses. The argument was that by the truth schema, it is both the case that '$2^{\aleph_0} = \aleph_1$' is true iff $2^{\aleph_0} = \aleph_1$ and that '$\neg(2^{\aleph_0} = \aleph_1)$' is true iff $\neg(2^{\aleph_0} = \aleph_1)$; so by excluded middle (and non-contradiction), either '$2^{\aleph_0} = \aleph_1$' is true or '$\neg(2^{\aleph_0} = \aleph_1)$' is true (and not both). But by eliminating excluded middle (and/or noncontradiction), the conclusion is avoided.

But is this enough to make positive sense out of indeterminacy? Intuitively, a sentence A is indeterminate if and only if there is no fact of the matter as to whether it's true. That is, iff it isn't a fact that it's true, and it isn't a fact that it's not true. That is, assuming that 'true' and 'it's a fact that' are the trivial operators the disquotationalist takes them to be, iff $\neg A \wedge \neg\neg A$. But presumably an advocate of one of these logics shouldn't count $\neg A \wedge \neg\neg A$ (or its equivalent $A \wedge \neg A$) acceptable, so we haven't made it acceptable to assert that there's no fact of the matter as to whether A is true. Nor can we achieve the effect by denying excluded middle: for that is equivalent to asserting $A \wedge \neg A$ in these logics.

One possible response is to deny the claim about the acceptability of asserting instances of $A \wedge \neg A$. The most plausible version of this would allow the "threshold of correct assertion" to vary with context: for many purposes it should be taken to be high enough to exclude anything of form $A \wedge \neg A$ and hence anything of form $\neg(A \vee \neg A)$, but for some purposes a lower threshold that includes some such sentences is acceptable. (By a threshold of correct assertion I mean, of course, a division of the "degrees of truth" into "good enough" and "not good enough", with the set of values that are "good enough" upward-closed under the partial ordering of the degrees.) When we say that there is no fact of the matter as to whether A, we are adopting a low enough threshold to include some sentences of form $\neg(B \vee \neg B)$ (and $(B \wedge \neg B)$), and are holding A to be among them. An attractive feature of this proposal is that it allows what is asserted in an indeterminacy claim to vary contextually, in a natural way: in some contexts the threshold may be set low enough to allow for lots of indeterminacy, while in others it may be set high enough to allow much less. (And if degrees aren't linearly ordered, thresholds needn't be either, so that in some contexts we can allow for lots of indeterminacy about baldness and little about richness, while in other contexts it can be the other way around.)

It isn't at all clear, though, that a deflationist should be happy about talk of "degrees of truth" (even if it is not assumed that the un-degreed notion of truth is definable from it). Nor is it clear that using a more neutral terminology like "semantic value" would help. It's hard to object to introducing a lattice of semantic values as an algebraic tool for determining what's a good inference and what isn't, but using it in an account of "correct assertion" as done in the last paragraph may well seem suspect.

Of course, the main idea of the last paragraph was that the standard of correct assertion should be allowed to vary contextually, and that in some contexts it is perfectly OK to deny excluded middle; and I'm inclined to think that this idea could survive even without the backing in terms of degrees of truth. But I will not pursue this; from here on out I will grant the assumption that instances of excluded middle are never to be denied.

Given this, we're back to our problem: we're in no position to accept of any sentence that there is no fact of the matter as to its truth.

But perhaps, though we can't *accept* the claim that there is *no* fact of the matter, we can *reject* the claim that there *is* a fact of the matter? This of course would require a notion of rejection that doesn't amount to acceptance of the negation. Can a suitable such notion be found?

There is an obvious *unsuitable* notion: refusal to accept. It does have one virtue: it does seem fairly natural to suppose that if we are convinced that a certain claim is indeterminate (e.g. the continuum hypothesis, or the claim that Joe is bald), then we should refuse to accept excluded middle for it. The problem is that we want to allow for uncertainty as to whether Joe is an indeterminate case of baldness, say when we've seen him only from a distance. Similarly, we want to allow for uncertainty as to whether the continuum hypothesis is indeterminate. (Maybe we don't know of the independence proof, or don't know of Putnam's argument that there is no way of giving sense to the idea of an "intended" model, or don't know whether there is a good reply to it.) And if we don't accept excluded middle for sentences we're convinced are indeterminate, we presumably shouldn't accept it for sentences we thinks are *very likely* indeterminate either; or even, for sentences for which we think there's about a 50% epistemic probability that they are indeterminate. But if the latter, then *unwillingness to accept* excluded middle is too weak to serve as a test of belief in indeterminacy.

What we need, then, is a sense of rejecting excluded middle that is stronger than refusing to accept it, but weaker than accepting its negation. Is there such a notion? I once thought it would be difficult to find one (Field 1994a), but now think I was missing the obvious. Let's think of the matter in terms of degrees of belief. Accepting a sentence is intimately connected with having a high degree of belief in it: say, higher than a certain threshold T, where $T>\frac{1}{2}$.[9] Similarly, I suggest, rejecting it seems connected with having a correspondingly low degree of belief in it: lower than the *co-threshold* 1-T, which is less than T. So rejecting it is certainly stronger than not accepting it: it requires a lower degree of belief. Is rejecting weaker than accepting the negation? *Given the assumption that degrees of belief obey the classical probability law $P(\neg B)=1-P(B)$*, rejecting A in the sense defined is precisely equivalent to accepting its negation. But it seems intuitively clear that if we abandon excluded middle we ought to modify the probability calculus to allow that $P(B)+P(\neg B)<1$, and this makes it possible to reject B while also rejecting $\neg B$.

In particular, we can do this when B is of form $A \vee \neg A$. It might have seemed impossible to make sense of assigning a degree of belief less than 1 to $A \vee \neg A$: how could that have probability 1 without its negation having probability greater than 0? And yet its negation is equivalent to $A \wedge \neg A$, so presumably *can't* have probability greater than 0. But on the nonclassical probability option, one can maintain that the probability of $\neg(A \vee \neg A)$ is always 0, while holding that $P(A \vee \neg A)$ is nonetheless less than 1.

One might try to avoid introducing nonstandard degrees of belief, and invoking instead the notion of truth. One often hears it said that although the proponent of a logic without excluded middle shouldn't deny any instances of excluded middle, she should deny of certain instances *that they are true*. (Take denial to mean simply, acceptance of negation.) If this were allowable, then rejection of an instance of excluded middle could simply be explained as denial *of its truth*. Similarly, we could coherently have a degree of belief of 0.5 in the claim that $A \vee \neg A$ is true and a degree of belief of 0.5 in the claim that it isn't true, without any nonstandardness in our degrees of belief. Unfortunately, this line is *not* allowable, on any notion of truth that obeys the standard equivalence between $Tr(\langle B \rangle)$ and B: that equivalence precludes denying the truth of $A \vee \neg A$ without denying $A \vee \neg A$ itself, and we're conceding here that instances of excluded middle are never to be denied. If we are going to make sense of the idea of rejecting there being a fact of the matter in terms of a logic that gives up excluded middle, I think we have no alternative but to invoke nonstandard degrees of belief.

The proposal for making sense of indeterminacy, then, involves both a non-classical logic in which excluded middle does not hold in general, and a corresponding non-classical probability theory for our degrees of belief, in which $P(A) + P(\neg A)$ can be less than 1.[10] One believes a sentence determinate to the degree that one believes $A \vee \neg A$. To be convinced that Joe is determinately bald or determinately not bald (even if one doesn't know which) is to believe this disjunction to degree close to 1; to be convinced that Joe is a borderline case is to believe it to degree close to 0; to be unsure is to have a degree of belief more in the middle. Similarly for beliefs about the determinacy of the continuum hypothesis.

As stated so far, the proposal doesn't literally allow that there be a proposition about which people disputing the determinacy of the continuum hypothesis (or of Joe's baldness) disagree; rather, the disagreement is in attitude, about what degrees of belief to have. But there is a natural extension of the proposal, on which we can interpret discussions of determinateness and indeterminateness at face value. The extension is to introduce a new determinateness operator G, governed by certain constraints on the nonclassical degrees of belief in sentences containing it. Actually it's simpler to take as basic an operator D, where DA means that it is determinately the case that A. The claim GA that A is determinate (i.e., that it is determinate *whether* A) is the claim that $DA \lor D\neg A$. The constraints I propose for the application of D to atomic sentences (or atomic formulas under an assignment to the variables) are:

$$P(DA) = P(A)$$

$$P(\neg DA) \in [P(\neg A), 1 - P(A)].$$

(If one doesn't want to allow for higher order indeterminacy, one can simplify the latter to: $P(\neg DA) = 1 - P(A)$.) And I propose that to obtain the probability of any other sentence (or formula under an assignment) in the language with D, we drive the D inward, by successively replacing

$D(A \land B)$ by $DA \land DB$; $D[\neg(A \land B)]$ by $D(\neg A) \lor D(\neg B)$

$D(A \lor B)$ by $DA \lor DB$; $D[\neg(A \lor B)]$ by $D(\neg A) \land D(\neg B)$

$D[\forall xA(x)]$ by $\forall xDA(x)$; $D[\neg\forall xA(x)]$ by $\exists xD(\neg A(x))$

$D[\exists xA(x)]$ by $\exists xDA(x)$; $D[\neg\exists xA(x)]$ by $\forall xD(\neg A(x))$

$D(\neg\neg A)$ by $D(A)$.

(These rules for D have a somewhat unappealing consequence, noted in Fine 1975: they prohibit "penumbral connections" among predicates, so that to call something determinately either green or blue commits us to its being either determinately green or determinately blue. Note though that the advocate of this approach can always introduce a new primitive predicates 'blue-to-green', not equivalent to the disjunction but related to it by such laws as that blue-to-green and not blue entail green. I'll suggest a slightly more flexible version of this in the next section.)

Now that we have the D operator, we can use it to explain a "stronger than disquotational" notion of truth: A is strongly true iff DA is true in the ordinary disquotational sense; or equivalently, iff A is disquotationally true and GA.[11] We met the notion of strong truth three paragraphs back (paragraph beginning 'One might try to avoid ...'): it is what is involved when one, while refusing to assert $\neg(A \vee \neg A)$, does assert that neither A nor $\neg$A is true. That paragraph may have seemed to suggest that the disquotationalist would have a problem accepting such a notion of truth. But what I really meant was (i) that she can't accept it as her basic notion of truth; (ii) that she can't accept it all without some kind of explanation; (iii) that she can't use the strong notion of truth in explaining indeterminacy unless she can explain the strong notion of truth without relying on the notion of indeterminacy. The present discussion by no means takes this back: rather, what I'm now proposing is

(a) that we explain the notion of determinateness in the way we explain most logical notions, in terms of its conceptual role;

(b) that the relevant conceptual role is specified not just by the logical laws governing the notion, which don't take us very far, but also by other constraints on degrees of belief;

(c) that once we have the notion of determinateness, we get the strong notion of truth as a byproduct, defined from D or G and disquotational truth.

The order of explanation is the opposite of the one rejected in the earlier paragraph.

I conclude this section with the observation that the disquotationalist has special reason to take seriously the use of a non-classical logic of something like the sort discussed here: in classical logic or even intuitionistic logic one cannot consistently accept the intersubstitutability of T(<B>) and B, for all B, because of the Liar paradox and the Curry paradox. That equivalence can however be maintained, in the logics here considered.[12]

6. A classical variant. In Section 3 I considered the militant disquotationalist's rejoinder to the model-theoretic argument against the determinacy of typical undecidable sentences of set theory (a rejoinder which, you'll recall, seems not so different in spirit from Putnam's own rejoinder). But I noted that militant disquotationalism seemed grossly implausible because it disallows *all* indeterminacy, even in the case of vagueness. So I promised three ways to make disquotationalism less implausible. The first was the one developed by Stephen Leeds. The second was the one just considered, involving non-classical logic. (Actually the second one subdivided, into a simple version without a determinateness operator and another version, more satisfactory I think, that uses the simple version to introduce a determinateness operator.) Now for the third.

Despite the significant reason given at the very end of the last Section for favoring a non-classical logic over classical–viz., the semantic paradoxes–classical logic has obvious virtues of its own. Is it possible to adopt the core idea of the non-classical logician's response to the militant disquotationalist, but within classical logic? It is indeed. I'll be sketchy about this, because I've developed the details elsewhere. (Field 2001, Ch 10 and the Postscript to it. The treatment there was for the classical case directly, rather than the idea here of basing it on the non-classical case.)

We've seen that a main component of the solution involving non-classical logic isn't the non-classical logic itself, but the associated theory of epistemic probabilities (degrees of belief). Mightn't it be possible to respond to the militant disquotationalist by adopting a non-standard probability theory even in the context of classical logic?

The probability theory won't be quite the same as the one for the non-classical logic, because any decent probability theory for classical logic must give all theorems of classical logic probability 1, and so in particular $P(A \vee \neg A)$ must always be 1. But we can still get that $P(A) + P(\neg A) < 1$ for certain A, if we give up the law (accepted in the non-classical case: see note 9) that $P(A \vee B) + P(A \wedge B) = P(A) + P(B)$. A neat probability-like theory based on an alteration of this law was developed in Shafer 1976, and it turns out that reflections rather like those in the previous section, but in a classical logic context, lead to it.

Recall that the solution in the last section had two stages. The first stage did not make it literally possible to believe that any claims were indeterminate: no notion of determinacy was introduced. What it did do was to give intelligible laws governing degrees of belief that allowed instances of excluded middle to have degree of belief less than 1. This allowed us to do two things: (I) to capture the idea of believing to a certain degree that a given claim is determinate; (II) to defuse a certain argument for the unintelligibility of indeterminacy. Now (II) is trickier in the classical logic case: we can't simply point out that the argument relied crucially on excluded middle. But part (I) works very much as it does for non-classical logic: by altering the laws for degrees of belief, we can use the extent to which $P(A) + P(\neg A)$ falls short of 1 as a measure of the extent to which one believes A indeterminate.

The second stage in the non-classical case was to use the non-standard degrees of belief to literally introduce a notion of determinateness into the language: the degrees of belief allowed us to give a conceptual role account of a determinateness operator G. This can be done too, in the classical case. As in the non-classical, we take as basic a simpler operator D, where DA means that it is determinately the case that A; GA, meaning that it is determinate whether A, is defined as $DA \vee D\neg A$. The laws governing D are very natural (they make it a necessity-like operator, and degrees of belief in DA and $\neg$DA for atomic A governed by the same rules as in the non-classical case); and we get the nice result that $P(GA)$ is $P(A)+P(\neg A)$. And here too the approach allows for higher order indeterminacy.

I think that this provides a very natural account of the conceptual role of our beliefs about determinacy and indeterminacy, in a classical logic setting. But a question remains: how exactly are we to defuse the argument against the coherence of indeterminacy (argument (B) of Section 3)? After all, the reply to that in the case of non-classical logic turned on the rejection of excluded middle.

One way of trying to reply to that argument is to use the idea of strong truth, according to which for A to be strongly true is for DA to be (weakly, or disquotationally) true; equivalently, for A and GA both to be (weakly) true. The idea would be that what we mean when we say "there is no fact of the matter as to whether A" is simply that neither A nor $\neg$A is strongly true, i.e. that neither DA nor D$\neg$A is true in the ordinary sense. This would of course be a totally cheap and unilluminating reply if we hadn't given a conceptual role account of the operator D, and thus of strong truth, in terms of our degrees of belief. But we have.

Even so, I don't think that this is by itself an adequate reply to Argument (B). The challenge that argument (B) poses is to explain why it is reasonable to say "there is no fact of the matter as to whether A", in situations where A holds but DA doesn't or where

$\neg$A holds but D$\neg$A doesn't. It may well be that what we mean when we say that there is no fact of the matter as to whether A is that neither A nor $\neg$A is strongly true; still, that doesn't answer the question of why by 'no fact of the matter' we *should* mean lack of strong truth as opposed to lack of weak truth. To put the matter more pointedly, the functional import of the notion of "no fact of the matter" is supposed to include this: for anyone who believes that there is no fact of the matter as to whether A, it would be pointless and misguided to wonder whether A, or to care whether A, or anything like that. But even given the above account of the conceptual role of D and the explanation of 'no fact of the matter' in terms of it, it isn't at all clear why there being no fact of the matter as to whether A *should* have that functional role. Why is it incoherent to say: I know that $\neg DA \wedge \neg D\neg A$, but nonetheless I wonder whether A, and I very much hope that $\neg$A? Given that by classical logic $A \vee \neg A$, the reason for the incoherence of this is far from obvious.

It's possible that a fuller account of the role of the nonstandard degrees of belief would answer this question. But I'm now tempted by a different kind of response, which is to say that even if we adhere to classical logic when dealing with indeterminate language, we should regard it as simply a convenient device that doesn't get us into trouble if we use it in a certain limited way, to be described below; the ultimately more basic logic is the sort of nonclassical logic discussed in the previous section, for which excluded middle fails.

I think that there is a way to make sense of this in terms of the nonstandard degrees of belief that I've advocated for both the classical logic and non-classical logic cases; the basic idea is that the degree of belief functions based on classical logic really just serve to encode a class of degree of belief functions based on nonclassical logic. Here's how the idea works for a propositional language L (i.e. one without quantifiers or the new conditional suggested in note 11). Let an *elementary sentence* be a conjunction of atomic sentences and their negations. I've pointed out elsewhere (Field 2001, Postscript to Chapter 10) that for each (nonstandard, i.e. Shafer) degree of belief function Q for classical logic there is a unique degree of belief function R_Q for Kleene logic *that gives exactly the same degrees of belief for elementary sentences*.[13] This suggests a radical idea: that one might take Q seriously only as regards its pronouncements about elementary sentences.

One reason this idea may seem unappealing is that it throws away a strong advantage of the classical approach to indeterminacy over the non-classical: the allowance of penumbral connections among predicates, so that something can be determinately either blue or green without being determinately blue or determinately green. To allow for this, I propose that we generalize the process of getting a non-classical function R from a classical function Q. The idea of the generalization is to allow the non-standard logic approach to treat as atomic certain sentences that aren't atomic. Suppose we have singled out some special class Σ of sentences of L to be "pseudo-atomic", for instances, disjunctions between predicates which intuitively have penumbral connections. (For instance, Σ might be the set of sentences that are either of form 'Blue(t) $\vee$ Green(t)' or of form 'Red(t) $\vee$ Orange(t)'.) Form a new language L_Σ whose atomic sentences are the atomic sentences of L together with the sentences $\lfloor A \rfloor$ for $A \in \Sigma$; here the special brackets '$\lfloor$' and '$\rfloor$' are a new notation in L_Σ. The sentences of L_Σ are just the truth functional combinations of this expanded class of atomic sentences. We extend the function Q on L to a function Q_Σ on L_Σ in a trivial way, by "making the brackets invisible": that is, we require that $Q_\Sigma(\lfloor A \rfloor \equiv A)$ always be 1 for all $A \in \Sigma$; this guarantees that for any sentence B of L_Σ, if B^- is the result of replacing any special brackets in it by ordinary parentheses then $Q(B) = Q(B^-)$. But if we now transition from Q_Σ to a Kleene-probability function $R_{Q\Sigma}$ by the method mentioned previously, the result is *not* blind to the brackets. In particular, if A and B are elementary sentences of L, then $R(\lfloor A \vee B \rfloor)$ is $Q(A \vee B)$ (since $\lfloor A \vee B \rfloor$ is elementary and the brackets are invisible to Q_Σ and Q_Σ extends Q); but $R(A \vee B)$ is $Q(A) + Q(B) - Q(A \wedge B)$ (since A, B and $A \wedge B$ are elementary). Consequently, $R(\lfloor A \vee B \rfloor) > R(A \vee B)$ whenever $Q(A \vee B) > Q(A) + Q(B) - Q(A \wedge B)$. So by "bracketing" the sentences in Σ, we have allowed the resulting Kleene-probability function to encode more information, viz. the penumbral connections included in Σ according to Q.

As a limiting case, one could bracket every sentence, thus in a certain sense building the complete information about Q into R. This wouldn't be totally trivial, in that for sentences A one took to be indeterminate, one would have $R(A \vee \neg A)$ and $R(\lfloor A \rfloor \vee \neg\lfloor A \rfloor)$ low (though $R(\lfloor A \vee \neg A \rfloor = 1)$; and that's all one needs to reject there being a fact of the matter as to whether A. It may be more in the spirit of the non-classical approach, though, to bracket only sentences that it would be natural from the agent's point of view to treat as atomic: 'blue-to-green' and 'red-to-orange' might be natural candidates for predicates to be used in such sentences. In that case, Q contains "excess information", from the point of view of a non-classical logician, though some of this excess information would become relevant were one to treat more sentences as atomic.

I can now restate the radical proposal in a more flexible form: the idea is that one should take Q seriously only as regards its pronouncements about *generalized* elementary sentences. More fully, the idea is that the use of excluded middle is simply a convenient device whose use is best justified by its not leading to any different degrees of belief in these generalized elementary sentences that we wouldn't get from the ultimately superior nonclassical logic. Of course, classical logic would be effectively correct for contexts where we believe we have determinacy, since for such contexts we'd have the extra premise of excluded middle (and the extra rule $A \wedge \neg A \vdash B$). But non-classical logic is basic, and this fact, together with the account of rejection developed in the previous section, allows us to make clear sense of our rejection of a fact of the matter in cases of indeterminacy.

7. Equivalent descriptions. As I said at the start, it is natural to interpret the term 'metaphysical realism' in such a way that one is abandoning metaphysical realism in mathematics if one holds that typical undecidable questions are indeterminate in truth value. I have sketched several ways in which the idea of such indeterminacy might be filled out. But the question arises whether the picture of indeterminacy that emerges from these sketches has much to do with metaphysical irrealism as Putnam understood it.

I'm inclined to think that the answer is yes. In my view, the most suggestive aspect of Putnam's discussion of metaphysical irrealism has been his doctrine of "equivalent descriptions", according to which prima facie incompatible theories often amount to the same thing. What I'd like to do in this final section is suggest that the kind of machinery suggested for making sense of indeterminacy can also be used to shed light on the phenomenon of equivalent descriptions.

Let's consider a very simple example of the sort of thing Putnam had in mind: not a very exciting example, but one that avoids some distracting issues that may arise with more interesting examples. The example, from Chapter 1 of Putnam 1987, involves two trivial theories. T_1 says that the universe contains n point particles and nothing else, and says something about their properties and locations. T_2 says that the universe contains 2^n-1 objects: viz., the mereological sums of n point particles, i.e. the n point particles together with the 2^n-n-1 mereological sums of two or more such objects. And it agrees with T_1 about the locations and properties of the n particles. (There is also a third theory T_3, that says that the universe contains 2^n objects, viz., those of T_2 plus an additional "null object".) Putnam's view–and it is hard not to sympathize with it–is that any debate between these theories is empty: there is no fact of the matter as to which is right.

Putnam argues, though, that it is not easy to make sense of the emptiness of the debate without giving up "metaphysical realism". And I think that even without being very precise about what "metaphysical realism" comes to, we can get a feel for what is bothering him.

The problem is *not* that advocates of one of these theories can't recognize a dispute with advocates of other of these theories as verbal. Certainly an advocate of T_2 can do this: she merely says that the advocate of T_1 is employing restricted quantifiers. "Although the advocate of T_1 appears to be saying that the only things that exist are the point particles, he doesn't really mean *exist* when he says 'exist', he means 'exists and isn't a sum of two or more objects'. Once this is recognized, we see that the advocate of T_1 isn't disagreeing with me." The advocate of T_1 may be able to do something similar, by interpreting the quantifiers of the advocate of T_2 as expressing some sort of "pretend existence". Or alternatively, maybe he interprets the T_2-theorists sentences at other than face-value: their real content is simply that part of their apparent content that concerns point particles. (There may be some difficulty making literal sense of one or both of these lines within T_1; let us not go into that.)

But even if advocates of either doctrine can recognize the distinction as verbal, this doesn't seem to do full justice to the emptiness of the dispute. For as Putnam says, either way of regarding the distinction as verbal is a *partisan* one, it assumes one of the two views and argues that anyone apparently advocating the opposing view can be interpreted as not really doing so. What Putnam wants is a way of making sense of the idleness of the dispute that isn't, in this sense, partisan.

Is there a way of avoiding partisanship? At one point Putnam seems to suggest that the key is to give up the erroneous supposition that 'exists' is univocal. But that doesn't really help. For the partisan of the broader ontology can grant that 'exists' isn't univocal: it could mean unrestricted existence, or it could mean the more restricted ("exists and isn't a non-trivial mereological sum"). Similarly, if the partisan of the narrower ontology can interpret the partisan of the broader ontology in terms of pretend existence, he can regard that as an alternative possible meaning of 'exists'. The partisanship comes out not in the insistence that 'exists' is univocal, but in the biased way of explaining alternative existence concepts.[14]

Can partisanship be avoided? It certainly can in one way: simply refusing to take a stand between the ontologies, i.e. being agnostic. But that doesn't give what Putnam wants either: he wants that there is simply no matter of fact to take a stand about. And this of course brings us to a familiar problem: how can there fail to be a fact of the matter? For, putting aside both genuinely distinct theories and additional theories like T_3 that raise the same problem, we have that either T_1 is true or T_2 is true but not both. But then, which is it?

The problem, actually, is in some ways worse than for vagueness and indeterminacy. One familiar approach to making sense of vagueness and indeterminacy is in terms of the theory of reference: nothing in our use of 'rich' could have settled whether we are taking those with such and such assets and liabilities to be included in the extension. That's basically the approach to indeterminacy that is assumed in the Putnamian argument I considered in Section 1. Putnam rejects that approach, as we saw in Section 2, and I think he's right to do so, but it is an approach that makes a certain amount of initial sense. But the approach seems to require a commonality of underlying ontology: the apparent issue as to whether Jones is rich is an issue about the application of 'rich' to entities that exist whichever stand we take on the breadth of 'rich'. It is quite unclear how to adapt this approach to disputes over the breadth of application of 'exist': at least, how to do so without taking a partisan approach in favor of the broader ontology.

Do any of the models for understanding indeterminacy developed in Sections 4-6 help? It doesn't seem that the Leeds' model does. For it's hard to see how there can be any reasonable interpretation of 'serious language' on which existential quantification won't be part of such language; so unlike apparently empty questions about richness, which simply can not be stated in "serious language" under some reasonable interpretations of that, apparently empty questions about ontology will surely be expressible.

But the models involving non-standard degrees of belief seem more promising. The attitude of thinking that either T_1 or T_2 is correct but that there's no fact of the matter as to which does seem to be fairly well captured in the model of Section 6, by attaching a high degree of belief to their disjunction but degree of belief 0 to each disjunct. And this combination of degrees of belief doesn't require that we in any way privilege one of the theories over the other. The attitude is, perhaps, captured even better on the alternative model in Section 5, in which classical logic is rejected. Here we assign a degree of belief 0 to the disjunction of T_1 and T_2, as well as to T_1 and T_2 themselves, so there is no case to be made that we regard one of the two as correct; at the same time, we assign a low degree of belief to the negations of these theories, and a high degree of belief to the negation of the disjunction of the alternatives to them, making clear the sense in which T_1 and T_2 unlike the alternatives to them are acceptable. I know of no better way to capture the attitude of believing that there is no fact of the matter as to which of T_1 and T_2 is correct.[15]

Notes

1. In a footnote added when the paper was republished he says: "I no longer (1974) agree that the notion of 'set' presupposes the notion of *definability* ..." [Putnam 1975, p 19]. This is puzzling, because the paragraph of the original paper to which the footnote was attached argued that the totality of definable subsets is perfectly clear, it is the totality of

undefinable subsets of an infinite set that is unclear.

2. Though I suspect that the time between the writing of the two articles was substantially longer than the publication dates would suggest.

3. Just as there is an issue as to whether second order quantification is determinate, there is an issue as to whether even first order quantification is completely determinate. (The most obvious possibility here is an indeterminacy as to whether the quantifier is restricted; this was alluded to above in connection with McGee.) There is even an issue as to whether truth functional operators are completely determinate. In taking the notion of model and the usual valuation procedure for models for granted, Putnam's model-theoretic arguments might be thought to *under*state the possibilities for indeterminacy. I will not pursue this worry here.

4. Parsons 1990 gives analogous arguments in the case of number theory. In my critique of Lavine, McGee, and Shapiro I asserted that the same critique applied in the number theory case, but I recently heard a lecture by Parsons in which he plausibly differentiates the cases.

5. Leeds is not explicit that the distinction between serious and non-serious is to be context-relative, but it seems to me that that is required if the proposal is to have a chance.

6. To say that the translations are incompatible may involve saying that w** is not synonymous with w*; but w* and w** are both in our language, so this would not require the *inter-linguistic* notion of synonymy here at issue.

7. There is a complication: to handle the applications of mathematics to nonmathematical domains like physics and psychology, we need that these set theories are not merely consistent but jointly consistent with our consistent nonmathematical theories. And we must also suppose that each theory T of physics or psychology or whatever formulated in our ordinary set theory could be reformulated using the alternative platonistic theory. One way to do so would be to replace it by a nominalistic theory. Another would be to replace it by the claim that all of the nominalistic consequences of T-plus-some-chosen-set-theory are true. This second route may seem like a cheap trick. Note though that on the second route, the chosen set theory needn't be one that decides the size of the continuum, though it could if that turned out to be useful in the particular physical or psychological theory in question. Also, the set theory chosen for one physical or psychological theory needn't be compatible with the one chosen for another: this makes clear that the truth of the set theory is not being assumed in the superior conceptual framework, only its instrumental utility in a particular application.

8. Given such a lattice, the rules of valuation for $\wedge$, $\vee$, $\neg$, $\forall$ and $\exists$ are completely obvious. My preferred notion of validity for a semantics based on such a lattice is: an inference is valid iff the degree of truth of the conclusion is at least as great as the degree of truth of each premise. (Degrees needn't be linearly ordered.) If we assume that the conditional is material (i.e., $A \supset B$ just means $\neg A \vee B$), the resulting logic is sometimes called FDE (e.g. in Priest 2001). It invalidates the rule $A \wedge \neg A \vdash B$ as well as invalidating excluded middle.

It turns out that the set of valid inferences of FDE is unaffected by restricting to the semantic values in the Dunn 4-lattice:

DIAGRAM

If a sentence has value u, so does its negation, and the same for v; whereas of course 0 and 1 flip. (This lattice invalidates the inference from $A \wedge \neg A$ to $B \vee \neg B$; no lattice with linearly ordered degrees will invalidate it.) Even so, more general deMorgan lattices are useful in connection with the pragmatics of assertion. They are also useful in connection with the logic of connectives not definable from $\wedge$, $\vee$, $\neg$, $\forall$ and $\exists$: for instance, non-material conditionals.

(It may be noted that I am implicitly rejecting the usual interpretation of the 4-lattice, in terms of "truth value gaps" and "truth value gluts". The reason is given in Field 2001, pp. 145-6.)

9. This threshold in degrees of belief is different from the threshold in degrees of truth mentioned earlier. (We have agreed to assume a fixed threshold of degrees of truth, high enough to exclude any sentences of form $A \wedge \neg A$. The assignment of probabilities is presumably best construed as relative to that threshold: viz., as the probability of having a degree of truth over that threshold. The threshold for acceptance is a further threshold, in how high the probability has to be for the sentence to be accepted.)

10. Probabilities are to be real numbers in the interval [0,1]; if A implies B in the logic, $P(A) \leq P(B)$; $P(A \vee B) + P(A \wedge B) = P(A) + P(B)$; and $P(A \wedge \neg A) = 0$. It follows that $P(A) + P(\neg A) \leq 1$.

11. A simple way to extend a lattice-theoretic semantics of the sort considered in note – to the language with D would be to take DA to have value 1 if A does, otherwise to have value 0. More generally, one could take DA to have value 1 if A was in some given filter on the lattice, otherwise to 0. Both of these proposals rule out any higher order indeterminacy, but there are subtler extensions that allow for that.

12. Admittedly, there is something unsettling about the solution, if one doesn't expand the logic to include a non-material conditional: for without it, sentences of form $A \Leftrightarrow A$ are not

in general assertable, and so the intersubstitutability of T(<B>) and B does not lead to the assertability of T(<B>) $\leftrightarrow$ B. While it is tricky to find a suitable conditional that is arbitrarily embeddable, without introducing further paradoxes, it is not impossible: indeed, a conditional that corresponds to a derivability predicate in an arithmetical theory for the $\Rightarrow$-free logic works. (See Field forthcoming.)

13. To get the value $R_Q(A)$, first put A into a Kleene-equivalent disjunctive normal form (being careful not to use excluded middle to simplify); take $R_Q(A)$ to be the sum of the Q-values of the disjuncts, minus the sum of Q-values of conjuncts of two distinct disjuncts, plus the sum of the Q-values of conjunctions of triples of distinct disjuncts, and so on.

14. How about simply adopting two existence concepts, without saying how they relate? That seems like taking a partisan stand for the broader ontology: unless one does something to make clear that one has in mind a fictionalist or non-face-value reading of the broader quantifier, it is hard to see why one isn't advocating the ontology of one's broadest quantifier (or the union of the ontologies of one's quantifiers, if there is no inclusion).

15. Thanks to Kit Fine and Stephen Schiffer for helpful comments on an earlier draft.

References

Devitt, Michael 1983. "Realism and the Renegade Putnam". *Nous* 17: 291-301.

Field, Hartry 1991. "Metalogic and Modality". *Philosophical Studies* 62, 1-22.

--------------- 1994. "Are Our Logical and Mathematical Concepts Highly Indeterminate?" *Midwest Studies in Philosophy* 19: 391-429.

--------------- 1994a. "Disquotational Truth and Factually Defective Discourse". *Philosophical Review* 103: 405-452. Reprinted in Field 2001.

--------------- 1998. "Which Undecidable Sentences Have Determinate Truth Values?". In H. Garth Dales and Gianluigi Oliveri, ed., *Truth in Mathematics*, Oxford University Press, pp. 291-310. Reprinted with new Postscript in Field 2001.

--------------- 2001. *Truth and the Absence of Fact.*

--------------- forthcoming. "Saving the Truth Schema from Paradox".

Fine, Kit 1975. "Vagueness, Truth and Logic". *Synthese* 30: 265-300.

Glymour, Clark 1982. "Conceptual Scheming, or Confessions of a Metaphysical Realist". *Synthese* 51: 169-80.

Horwich, Paul 1998. *Meaning*. Oxford: Oxford University Press.

Lavine, Shaughan 1994. *Understanding the Infinite.* Cambridge, Mass.: Harvard University Press.

---------------------- (unpublished). *Skolem Was Wrong.*

Leeds, Stephen 1997. "Incommensurability and vagueness". *Nous* 31: 385-407.

Lewis, David (1984). "Putnam's Paradox". *Australasian Journal of Philosophy* 62: 221-236.

McGee, Vann 1997. "How We Learn Mathematical Language". *Philosophical Review* 106: 35-68.

----------------- 2000. "Everything". In G. Sher and R. Tieszen, eds., *Between Logic and Intuition: Essays in Honor of Charles Parsons.* Cambridge: Cambridge University.

Parsons, Charles 1990. "The Uniqueness of the Natural Numbers", *Iyyun, A Jerusalem Philosophical Quarterly* 39: 13-44.

Priest, Graham 2001. *An Introduction to Non-classical Logic*. Cambridge: Cambridge University Press.

Putnam, Hilary 1967a. "The Thesis that Mathematics is Logic". In R. Shoenman, ed., *Bertrand Russell, Philosopher of the Century* (London: Allen & Unwin). Reprinted in Putnam 1975.

----------------- 1967b. "Mathematics Without Foundations". *Journal of Philosophy* 64: 5-22. Page references to reprinting in Putnam 1975.

----------------- 1975. *Mathematics, Matter and Method: Philosophical Papers vol. 1*. Cambridge: Cambridge University Press.

----------------- 1978. "Reference and Understanding". In Putnam, *Meaning and the Moral Sciences*. London: Routledge.

----------------- 1980. "Models and reality". *Journal of Symbolic Logic* 45: 464-82. Page references to reprinting in Putnam 1983a.

----------------- 1981. *Reason Truth and History.* Cambridge: Cambridge University.

----------------- 1983. "Vagueness and Alternative Logic". In Putnam 1983a.

----------------- 1983a. *Realism and Reason: Philosophical Papers vol. 3*. Cambridge: Cambridge University Press.

----------------- 1987. *The Many Faces of Realism*. La Salle: Open Court.

Shafer, Glen 1976. *A mathematical theory of evidence.* Princeton: Princeton University.

Shapiro, Stewart 1991. *Foundations Without Foundationalism.* Oxford: Oxford University Press.

Weston, Thomas 1976. "Kreisel, the Continuum Hypothesis, and Second Order Set Theory", *Journal of Philosophical Logic* 5: 281-98.

Williamson 1994.