# *Weaseling Away the Indispensability Argument*

JOSEPH MELIA

According to the indispensability argument, the fact that we quantify over numbers, sets and functions in our best scientific theories gives us reason for believing that such objects exist. I examine a strategy to dispense with such quantification by simply replacing any given platonistic theory by the set of sentences in the nominalist vocabulary it logically entails. I argue that, as a strategy, this response fails: for there is no guarantee that the nominalist content of the platonistic theory is exhausted by this set of sentences. Indeed, there are platonistic theories that have consequences for the nominalist world that go beyond the set of sentences in the nominalist language such theories entail. However, I argue that what such theories show is that mathematics can enable us to express possibilities about the concrete world that may not be expressible in nominalistically acceptable language. While I grant that this may make quantification over abstracta indispensable, I deny that such indispensability is a reason for accepting them into our ontology. I urge that the nominalist should be allowed to quantify over abstracta whilst denying their existence and I explain how this apparently contradictory practice (a practice I call "weaseling") is in fact coherent, unproblematic and rational. Finally, I examine the view that platonistic theories are simpler or more attractive than their nominalistic reformulations, and thus that abstracta ought to be accepted into our ontology for the same sorts of reasons as other theoretical objects. I argue that, at least in the case of numbers, functions and sets, such arguments misunderstand the kind of simplicity and attractiveness we seek.

## *1. The indispensability argument*

The indispensability argument comes in two main flavours:

(1) We ought to believe the claims our best scientific theories make about the world—after all, they are our *best* scientific theories. But a casual glance through any book of theoretical physics reveals that our best scientific theories entail the existence of numbers, sets and functions: "The force between two massive objects is *proportional* to the *product* of the masses *divided* by the square of the distance" and "There is a one-to-one differentiable *function* from the points of space-time onto *quadruples of real numbers*" are typical of the kinds of sentences asserted by scientists and found in physics books. Since such claims entail the existence of *abstracta*, we cannot consistently assert or believe in our scientific theories whilst denying the existence of *abstracta*.[1]

[1] See Putnam 1971 for a vigorous presentation of this version of the Indispensability argument.

 

(2) We ought to believe in *abstracta* for the very same theoretical reasons we believe in the concrete, unobservable entities postulated by scientific theory. We postulate such things as quarks and space-time points not because we directly observe these entities, but for pragmatic or aesthetic reasons: because doing so either increases the explanatory power of our theory, or increases the theory's simplicity, or increases the theory's strength—or a combination of all three.[2]

The main aim of this paper is to show that (1) fails. I examine whether we can purge mathematics from science by replacing any given scientific theory with the set of sentences it entails that make no reference to abstract objects. Although I am sympathetic to the thought that the nominalist need only be committed to the nominalistic consequences of a given theory, I argue in §2–3 that this method does not necessarily give us a theory's nominalistic consequences and that another strategy is needed. Accordingly, I tackle argument (1) in a radical manner: in §5–6, I accept that we cannot formulate our scientific theories without quantifying over abstract objects, but I argue that, despite appearances, it is *not* contradictory or incoherent to quantify over abstract objects in one's scientific theory, and then go on to deny that there are any abstract objects. Sometimes it is legitimate to assert a collection of sentences whilst denying some of the logical consequences of this collection! Finally, in §7, I turn my attention to (2). Here I argue that abstract entities and theoretical entities are *not* on a par, that the reasons we postulate *abstracta* are quite different from the reasons we postulate quarks, electrons and space-time points and that the pragmatic reasons which may make it reasonable to accept quarks and space-time points do not apply to *abstracta.*

## 2. *Is mathematics really indispensable?*

As presently formulated, scientific theories do indeed quantify over abstract objects. So, if these theories are true, there are abstract objects. Nominalists cannot dismiss any scientific theory T quantifying over *abstracta* out of hand—for T may make claims about the concrete world which the nominalist cannot afford to ignore. Indeed, T may even have consequences about the concrete world which the nominalist has excellent reason to believe. "Ahh," we nominalists think to ourselves (for I am one), "if only there were a theory which, whilst not quantifying over abstracta, made precisely the same claims about the concrete world as T, which had the same consequences for the concrete part of the world as T,

[2] This line of thought can be found in much of Quine's work.

and which made the same predictions as T, then we would have a theory we could believe in which had all the advantages of T and none of its terrible disadvantages".

Some have argued that, even were we to succeed in reformulating our scientific theories so that they did not quantify over *abstracta* (as Field (1980) has tried to do), platonism is still the preferable position. After all, if and when a philosopher manages to find a reformulation of General Relativity which does not entail the existence of *abstracta*, the practice of physicists is unlikely to change. Physicists are overwhelmingly likely to keep using the heavy abstract machinery with which they have grown familiar, and keep on quantifying over *abstracta*. Accordingly, if we really are to respect the practice of scientists, we ought to accept the abstract ontology which their practice commits them to.

I do not think philosophers should bother themselves with this line of thought. It is quite common for both scientists and mathematicians to think that their everyday, working theories are only partially true. Instrumentalists may very well use a theory U whilst believing only in that set of sentences entailed by T whose vocabulary is taken solely from the observation language.[3] Finitists may very well use Peano Arithmetic whilst believing only in the set of sentences entailed by Peano Arithmetic whose quantifiers are all bounded. Perhaps some would regard such people as hypocritical, taking back by night what they practise by day. But I see nothing wrong in such an attitude. For if instrumentalists, finitists and nominalists have theories T committed to only kosher entities, and these theories have the same kosher consequences as the unkosher theories, then they can maintain that, though by day they may *use* unkosher theory U, what they really *believe* is theory T. True, the ontologically parsimonious must tell us why it is legitimate to use U when all they really believe is T (and, of course, it cannot be U's *truth* that makes it legitimate to use U). But I see no reason to think that *truth* is the only concept to which one can appeal in order to justify the use of a false theory (see, for instance, Field 1984).

There *is* a threatening charge of hypocrisy to be made, but the charge of hypocrisy is levelled at those who have no alternative kosher theory T to hand which expresses what they really believe. Without such a theory, the only option seems to be either to put up or to shut up: if you accept the theory then you must accept its unkosher consequences; if you cannot accept its unkosher consequences then you cannot accept the theory. You cannot have it both ways.

[3] Though, as we shall see later, there are some problems with this particular choice.

But could we ever really be in the position where we have an unkosher theory U but no kosher alternative T? For given a theory U it is surely a trivial matter to formulate an alternative theory T which has all the kosher consequences of U yet none of the unkosher ones. Given any unkosher theory U simply partition the predicates into two classes: those that are nominalistically acceptable and those that are not. Let theory T be those sentences that are logically entailed by U, yet whose vocabulary contains only nominalistically acceptable predicates. Since every nominalistically acceptable sentence logically entailed by U is logically entailed by T, it would seem as if T has all the kosher consequences of U and none of the unkosher ones, and thus can serve as the nominalist's replacement theory.[4] Since the reasoning here is entirely general, the nominalist appears to have a winning strategy no matter what the platonistic theory U that scientists eventually settle on will be.

I shall call this method of eliminating quantification over abstracta the *Trivial Strategy*. The trivial strategy certainly has all the advantages of theft over honest toil—but that does not necessarily mean that it has any of the disadvantages! Early on in his project to eliminate mathematics from science, Field does indeed consider such a cheap strategy (1980, p. 8). He rejects it, however, since he does not think that the theories in which it results are *attractive*. I shall question the validity of rejecting such unattractive theories in §7 but, even if he is right, in a very short time we have come a long way from the view that quantification over abstracta is *indispensable*. Quantification over abstracta can be dispensed with—and easily dispensed with at that—but the theories which do quantify over abstracta are more attractive than the theories which don't. This is a considerably weaker claim and one much more vulnerable to a nominalist assault.

In fact, there is no need to go into questions of attractiveness: in its current formulation the trivial strategy does not always work. For the following two possibilities can both occur: (1) a platonist theory can have certain consequences for the concrete world; (2) the set of all sentences restricted to the nominalist vocabulary that the theory entails does not have the same nominalistic consequences as the platonist one.

[4] This procedure is more familiar from the philosophy of science, where it was used by instrumentalists to get rid of unwanted ontological commitment to theoretical entities.

## 3. *Two theories of mereology*

In this section, I introduce two familiar first-order theories of mereology. Whilst T is a nominalist theory, T* is an extension of T which quantifies over functions and sets. It will turn out that T* is a conservative extension of T: T* entails no new sentences in the vocabulary of T. Yet T* does entail the existence of a kind of region which T does not guarantee to exist.

Essentially, T is nothing more than a simple first-order theory of mereology that also says there is no atomless gunk and that implies there are infinitely many atoms.

More formally: T contains the two place predicate $x < y$ – which means $x$ is a part (though not necessarily a proper part) of $y$. T also contains the usual axioms guaranteeing that $<$ is a partial order.

$$\forall xyz(x < y \,\&\, y < z \rightarrow x < z)$$
$$\forall xy(x < y \,\&\, y < x \rightarrow x = y)$$

etc.

It also contains a constant 1, with the axiom

$$\forall x(x < 1).$$

1 is the maximal element—the mereological sum of everything. It also contains the predicate Atom($x$)—to be read "$x$ is an atom".

In our theory, we will want Atom($x$) to be true of $x$ precisely when the only thing contained in $x$ is $x$ itself. So we include

$$\text{Atom}(x) \leftrightarrow \forall y(y < x \leftrightarrow y = x).$$

According to T there is no atomless gunk in our universe. Everything that exists is composed of atoms. Accordingly, T includes the axiom

$$\forall x \exists y(\text{Atom}(y) \,\&\, y < x).$$

According to T there are an infinite number of atoms, and so T contains the axioms:

$$\exists xy(\text{Atom}(x) \,\&\, \text{Atom}(y) \,\&\, \neg(x = y))$$
$$\exists xyz(\text{Atom}(x) \,\&\, \text{Atom}(y) \,\&\, \text{Atom}(z) \,\&\, \neg(x = y) \,\&\, \neg(x = z) \,\&\, \neg(y = z))$$

and so on …

T also has to say something about what mereological sums exist. Since T is a first-order language, it must do this by containing a comprehension schema. Although this is similar to the comprehension schema from set-theory, there are a couple of differences: (a) there's no empty region; (b) there isn't always a region containing *precisely* the things which are $F$ (i.e. there is not always a region containing all the $F$s and no things which are not $F$). Example: The mereological sum of three green mice contains parts which are not themselves green mice—such as one of their tails.

Informally, the comprehension schema says that, for any predicate $F$ of our first-order language, if there is something which satisfies that predicate then there is a region R which contains all the things which satisfy $F$ AND anything which shares a part with R shares a part with one of the things which $F$s.

More formally, for every predicate $F$,

$$\exists y(F(y)) \rightarrow \exists \mathrm{R}\{\forall y(Fy \rightarrow y < \mathrm{R}) \ \& \ \forall y(\exists z(z < \mathrm{R} \ \& \ z < y) \rightarrow \exists w(Fw \ \& \ \exists z(z < w \ \& \ z < y)))\}.$$

This ends the formal presentation of T.

Though T is not finitistic (since it postulates the existence of an infinite number of atoms) it is nominalistic. Since the atoms postulated by the theory may be taken to be concrete entities, and since mereological sums of concrete entities are themselves concrete, no commitment to abstract objects is entailed by the theory.

Our platonistic extension of T is also relatively straightforward. T* simply extends T in the following five ways:

(i) T* contains all the axioms of ZF set-theory.

(ii) T* also contains the axioms $\exists x(\mathrm{Atom}(x) \ \& \ \neg Sx)$; $\exists x \exists y(\mathrm{Atom}(x) \ \& \ \mathrm{Atom}(y) \ \& \ \neg Sx \ \& \neg Sy)$, (where "$Sx$" means "$x$ is a set")—and so on. In this way it says that there are infinitely many non-mathematical atoms.

(iii) T* includes all the axioms of T.

(iv) T* contains all instances of the mereological comprehension schema which include the vocabulary of set-theory.

(v) T* contains an axiom which asserts that the non-mathematical objects form a set.

Let us say that a region R is infinite and coinfinite iff R contains an infinite number of atoms and R's complement contains an infinite number of atoms.

Now, T has a model of the following kind. The model contains a denumerable number of atoms. There are an infinite number of regions, but every region either contains only a finite number of atoms, or is such that its complement contains a finite number of atoms (for proof, see the appendix). Since T permits such models, the theory cannot be said to entail the existence of a region which is both infinite and coinfinite.

T*, however, does not permit the existence of such models for this theory *does* entail the existence of regions which are infinite and coinfinite. For since the non-mathematical objects form a set, T* entails the existence of a one-to-one function F from this set into the ordinals. Then by the comprehension axiom of T*, there is a region containing precisely those atoms which are the range of the even ordinals smaller than $\omega$. Clearly such a region and its complement are both infinite. Accordingly, T* entails the existence of a kind of region which T does not. But T* does not entail any new sentence in the nominalist vocabulary of T. For T is a complete theory (for proof, see the appendix): for every sentence $\phi$ in the vocabulary of T, either $\phi$ or $\neg\phi$ is entailed by T. Since T* is clearly consistent, it follows that T* is a conservative extension of T.

Moreover, since T is complete, and since T* contains T, T *just is* the set of sentences entailed by T* in the nominalist vocabulary.

This result means that the nominalist cannot in general adopt the trivialisation strategy as outlined above. For I take the above to show that T* has implications for the nominalist world which T does not—even though T *just is* the set of sentences entailed by T* in the nominalistic vocabulary of T*. For if T* is true, then a certain kind of region exists: a region which contains infinitely many atoms and which fails to contain infinitely many atoms. Such a region is not guaranteed to exist by T alone. Since this is a fact about what kinds of *regions* there are, I take this to be one of the *nominalist* consequences of T*. Accordingly, it turns out that T* may have nominalistic consequences without there being any nominalistically acceptable sentence in the language of T* implied by T*. In other words, there can be more to the nominalist consequences of a theory than the set of sentences entailed by that theory in the nominalist vocabulary. If the nominalist simply takes his theory to be the set of nominalistically acceptable sentences entailed by some platonist theory, he has no guarantee that his theory actually has the same nominalist content as the platonist theory.

At first sight, this conclusion may seem paradoxical. How can the sentences T* logically entails in the nominalist vocabulary *not* have the very same consequences for the nominalist world as T* itself? But this air of paradox is illusory: there is simply no reason to think that the implications T* has for the concrete world are *ipso facto* expressible in the nominalistically acceptable vocabulary of T*. Above, when I argued that T* has new consequences for the nominalist world, I did *not* show that there was some nominalistically acceptable sentence entailed by T* but not by T, nor by showing that there was a nominalistically acceptable sentence that only T* could express. Instead, I compared the *models* of the two theories and noted that there were models of T which failed to contain a certain kind of region that appeared in every model for T*; one cannot immediately move from the fact that the models of the theories differ in this way to the conclusion that there is some nominalistically acceptable sentence one could add to T that would give us precisely the models the nominalist desired.[5] In particular, the way in which the theory T* ensures the existence of an infinite coinfinite region is by containing a sentence which asserts the existence of a sum whose atoms are mapped onto the even numbers by some 1-1 function from atoms to ordinals. Although this sentence has implications about what kinds of concrete regions there must be, it is not a nominalistically acceptable sentence.

Although I say that the result spells trouble for the trivialisation strategy it does not, of course, destroy the nominalist. For the trivialisation strategy is not the only strategy open to him. For instance, the nominalist may give up on T as a vehicle for expressing the nominalistic implications of T*, and might try to develop another theory T** which, on top of the axioms of T also contained the sentence "There exists a region containing an infinite number of atoms and whose complement contained an infinite number of atoms". But this is quite a different strategy. Whereas the trivialisation strategy restricted itself to the nominalistically acceptable sentences of the problematic platonic theory, this new strategy asks us to develop logical and semantic resources that go beyond it. I am sceptical of the prospects for this strategy. Such a strategy means that the nominalist must not only provide us with a new notion of the nominalistic content of a theory—it can no longer be simply the set of sentences entailed by the platonistic theory in the nominalist vocabulary—he must also convince us that whatever his new primitives may be, they are primitives which are acceptable to the nominalist.[6] This is not always a trivial matter. True, there is some plausibility in the view that "there are infinitely many" may be taken as a primitive, but there are also those who think that our only real grasp on the finite/infinite distinction *is* via the notion of one-to-one correspondences and set-theory. And it is moot whether this new strategy can successfully be extended to more complicated theories. For instance, a platonistic theory could introduce a predicate for those regions of space-time that corresponded to measurable sets of $R^4$, and could then go on to say something about the way these regions are. Since the regions themselves are nominalist entities, the theory would then have represented the nominalist world as being a certain way, even though it had used a lot of complicated platonistic apparatus to pick out these nominalist entities. I think it is dubious, to say the least, that there is a nominalistically kosher way of picking out such regions. Similarly, a platonist theory might contain the two-place predicate "*l* is root two as long as *m*". Such a sentence

[5] For instance, a nominalist theory can have two non-isomorphic models which are elementarily equivalent. The two models represent different ways in which the nominalist world could be—but there is no sentence in the vocabulary of the nominalist theory that is true in one model but false in the other.

[6] Field himself has some qualms about introducing such primitive cardinality quantifiers. See Field 1980, Ch. 9 and Field 1985.

has consequences for the concrete world, yet the sentence itself is not nominalistically acceptable: the number root two is mentioned. Again, it is implausible to take such a predicate as primitive and thus the nominalist must find some nominalistically acceptable way of expressing this relation.[7]

However, although the *letter* of the trivialisation strategy fails, I am still sympathetic to its spirit. The problem that emerges from the preceding discussion is this: how can the nominalist help himself to the nominalist consequences of a theory if those consequences are not necessarily expressible in a purely nominalistic language. In §5, I put forward a way in which the nominalist might try to do this. The method will require us to tackle the indispensability argument head on.

## 4. Three interesting implications

Before attacking the Indispensability argument I would like to point out three interesting philosophical implications of the above result.

1. It follows from the properties of T and T* that a body of mathematics can be added conservatively to a nominalistic theory, and the resulting theory can have consequences for the nominalist world which the initial theory does not. For T* just is T plus a body of mathematics. Yet, as we have seen, T* entails no new sentence in the nominalist vocabulary.

Field has conceded that mathematical entities do have some theoretical utility, but he argues that their theoretical utility is quite unlike the theoretical utility of concrete entities (Field 1980, Ch.1). He argues that mathematics is *conservative* over theories which are formulated nominalistically, and that *therefore* they add nothing new to the nominalistic theory. The above result shows that this simply does not follow. T* is indeed conservative over T, but T* has implications for the nominalist part of the world which T simply does not have.

2. This example also strengthens a point of van Fraasen's. Van Fraasen has pointed out that one cannot always identify the empirical content of a theory with the set of sentences entailed by that theory in some privileged

[7] Field does try to show how to talk about ratios in a nominalistically acceptable way—but his attempt, which involves postulating infinitely many points of spacetime and certain primitive relations upon these points, is clearly a world away from the trivialisation strategy. Indeed, it is not clear whether he has succeeded in formulating a theory capable of expressing all the facts about the *nominalist* world that typical platonist theories can. For instance, whilst he succeeds in showing how the nominalist may say that *rational* ratios obtain between lines, he does not show that all *irrational* ratios can be asserted by the nominalist. See Melia 1998 for further discussion.

vocabulary (van Fraasen 1980, pp. 54–5). For instance, it is a consequence of quantum mechanics that there exists an $x$ which does not have a precise location in space and time. The words of this sentence are drawn wholly from the observational vocabulary, but this is scarcely a sentence which the instrumentalist would accept. Van Fraasen has shown that there can be sentences in the kosher vocabulary which are nevertheless unkosher—that the set of sentences in the kosher vocabulary entailed by the unkosher theory can sometimes say too much—and van Fraasen uses this to motivate the idea that the empirical content of a theory should be identified with certain *submodels* of a theory rather than any special set of sentences entailed by the theory. Nevertheless, as van Fraasen is aware, there is a lacuna here for there still remains the possibility that there exists *some* set of sentences of the original theory, sets omitting the unkosher sentences, which do succeed in capturing its empirical content.

However, our two mereological theories show that sometimes the set of sentences in the kosher vocabulary *fails* to capture all the kosher consequences of the unkosher theory: the set of sentences says too little. But if the set of *all* sentences in the kosher vocabulary is too weak to capture the kosher consequences of the theory, then *no* set of sentences in the kosher vocabulary will capture all the kosher consequences. There is no subset of sentences of the unkosher theory capturing precisely the kosher content of the theory.

3. Finally, the example spells trouble for a formalist philosophy of applied mathematics. If mathematics were used only to simplify or permit deductions between nominalistically acceptable sentences, then a formalist view of mathematics would still be a player. But I do not see how the formalist can make sense of the above example. To see this, we must look in a little detail at just how T* manages to have implications for the nominalist world which T does not.

T says very little about the atoms themselves. Indeed, all it really says about these objects is that they are simple and that there are an infinite number of them. Since the atoms have so little structure it is hard to define predicates which pick out particularly interesting collections of atoms. In particular, it is not possible to define a predicate satisfied by an infinite coinfinite region.

By contrast, the natural numbers are a lot more complex and instantiate a richer structure. Not only are there an infinite number of them, but certain couples fall under the successor relation, others fall under the *larger than* relation. Certain triples fall under the addition relation, others fall under the multiplication relation. Similarly, the theory of Peano Arithmetic is richer than the simple theory of mereology developed above. Because the numbers have such complex structure and because the language of arithmetic is rich enough to talk about this structure, more complex sets of numbers are definable. In particular, it is possible to define the predicate "$x$ is even" in PA—and of course the set of numbers which satisfies this predicate is infinite coinfinite.

However, simply adding number theory to T will not immediately yield anything new about the regions. But adding number theory to T plus a one-to-one function f from the numbers into the atoms *will*. For once we have such a function, any predicate $P(x)$ of PA applying to a set of numbers corresponds to a predicate $P^*(x)$ applying to a set of atoms. $P^*(x)$ is simply: "there is a number $y$ with $P(y)$ and $\mathrm{f}y = x$". If $P(x)$ is "$x$ is even" then $P^*(x)$ picks out a collection of atoms which is infinite coinfinite. Now, providing our richer theory contains the predicate $P^*(x)$, the comprehension schema for regions guarantees the existence of a mereological sum which is infinite coinfinite.

The regions that are guaranteed to exist by such a comprehension schema depend upon two things: (1) the predicates contained in theory T; (2) the properties and relations T says are instantiated by the atoms. By adding extra predicates and extra ontology, we are able to generate new predicates which are satisfied by sets of atoms not definable in T alone. Just adding extra predicates or just adding extra entities and properties will not suffice. For instance, one could have simply added the predicates "$x$ is a set" and "$x$ is a member of $y$" to the language of T, and thus generate new instances of the comprehension schema of T—but without adding the extra ontology the resulting theory will not entail anything new about the regions. Similarly, one could simply add extra ontology and postulate the existence of the set-theoretic hierarchy. But without augmenting T with extra linguistic resources (the predicates "$x$ is a set" and "$x$ is a member of $y$") one will not generate any new instances of the comprehension schema, and the resulting theory will not entail anything new about what kinds of regions and sums there are.

Now we can see that the above explanation of how the mathematics is applied is not consistent with a formalist theory of applied mathematics. The mathematics is not facilitating or enabling us to make inferences between (meaningful) nominalistic sentences. Indeed, the mathematics does not give us any new nominalist sentences. Nor will we get any new results about regions by merely adding meaningless symbolism. Defining new regions demands taking the mathematics seriously: unless there really are the numbers over and above the atoms, instantiating new properties and relations, no new regions will be definable. For the mathematics to be applied, it must be regarded as meaningful.

But although the mathematics cannot be regarded as meaningless symbolism, it does not matter exactly what we take the predicates "$x$ is a set" and "*x* is a member of *y*" to mean. Any interpretation of the predicates will do, providing that under this interpretation the axioms of set-theory are still true. So a structuralist interpretation of mathematics is certainly possible. It does not matter whether or not we take sets to be located, whether or not we take them to have causal powers, whether or not they are coloured—as long as there are entities playing the right kind of role, we have the wherewithal to talk about infinite coinfinite regions.

## *5. Against the indispensability argument (1): the way of the weasel*

I now return to the indispensability argument. The theories T and T* showed us that one cannot always get around this argument by taking all the nominalist consequences of a platonistic theory. For the platonist language enabled us to say something about the way the sums were that was not expressible in the nominalist language. By adding new predicates to our theory and new entities to our ontology, we were able to guarantee the existence of a certain kind of concrete entity whose existence we were unable to guarantee in the nominalist system. Now, some nominalist N may well come to believe that T is wrong about the way the sums are: it does not entail the existence of an infinite coinfinite region. In this respect, N prefers T*, for T* gets it right about what kind of sums there are. Accordingly, since N wants a theory which captures his views about what regions there are, it seems as though N must use the mathematics of T* in expressing his views about what sums there are, and that he should accept T*. But how can a nominalist consistently accept T*? Surely it is simply inconsistent for N to use mathematical entities, assert sentences which entail that these entities exist, and then turn around and deny that there are such things as these entities. How can anyone coherently assert P, know that P entails Q, yet deny that Q is the case? How could it ever be rational to assert that P whilst denying a logical consequence of P?

Consider some hypothetical schmoe: Joe. Joe believes that T does not capture the whole truth about the sums and the atoms. He thinks there are more sums than T says there are. Joe prefers T*. For T* implies the existence of a kind of sum which T does not, and Joe thinks that such sums do indeed exist. However, Joe is no platonist. He realises that T* entails the existence of a certain kind of entity which Joe does not believe in: sets. So although T* is right about the sums, it is wrong about the platonistic entities. In order to communicate his picture of the world, Joe must make clear that, as far as he is concerned, T* is wrong in certain respects. What Joe wants to do is *subtract* or *prune away* the platonistic entities whose existence is entailed by T*. Accordingly, in communicating his belief about the world, Joe might say "T*—but there are no such things as sets". In doing so, he is asserting some sentences, whilst denying one of their logical consequences.

I emphasise that Joe does *not* simultaneously hold contradictory beliefs. Just because, in the process of telling us his beliefs about the world, Joe asserts all the sentences of T*, it does not follow that Joe believes all the sentences of T*. Indeed, since Joe believes there are no abstract objects, he will explicitly say that T* is false. So why is Joe asserting T* if he doesn't believe it? Because the mathematical structure in T* allows Joe to express possibilities for the *concrete* part of the world—possibilities that are not expressible without that structure. Joe is taking advantage of the mathematics in T* to *communicate* or *express* his picture of what the world is really like.[8]

In general, we assert sentences in order to present a picture of the way we think the world is. We normally think of each successive sentence in our story as adding a further layer of detail, either making explicit what was only implicit before, or filling in gaps and adding details not filled in before. With each sentence, our picture of the world becomes a little more filled in, the story we are telling a little more determinate. But *must* our stories about the world necessarily take this form? *Must* we think of each successive sentence as adding a layer of detail, or filling in the gaps which were left by our previous sentences? Why can't we understand some of the later sentences as *taking back* things that were said earlier on in the story? Why can't we understand the later sentences as *erasing* some of the details implied by the earlier part of the story? or as *changing* some of the implications of the previous parts of the story. That's what Joe is doing when, after having asserted T*, he goes on to add "but there are no such things as sets". He's just explaining that the implications T* has for the existence of abstracta are in fact not true. But the rest of T*, the implications T* has for the way the sums are, this much of T* he believes.

Taking back or tailoring part of what we have said before is not unknown in our everyday practice. "Everybody who *F*s also *G*s. Except Harry—he's the one exception". These are certainly sentences of English and it is clear what is being said. First, we assert that all the *F*s are also *G*. Then we go on to retract an implication of this sentence. We have said something which entailed that, if Harry *F*-ed then he also *G*-ed but, as a matter of fact, this is not so: Harry both *F*s and doesn't *G*. Have we

[8] It is often said that mathematics is a language. Of course, strictly speaking, this is not true. But we here see what is right about this idea. At least in the examples under consideration here, mathematical objects are postulated in order to make certain thoughts or beliefs expressible or communicable.

spouted contradictory nonsense? Of course not: what we have said is perfectly clear: apart from Harry, anyone who *F*s also *G*s. Of course, in this example, we *could* have simply said: "for any $x$, if $x$ is not identical to Harry then, if $Fx$ then $Gx$", and so say what we want to say without ever having to take back anything we ever said. But even so, "Everybody who *F*s also *G*s. Except Harry" makes perfect sense. We know exactly the thought being expressed by these sentences whether or not the neat formulation in first-order predicate calculus ever occurs to us.

Taking back part of what we have said before is not unknown in our scientific practice. We sometimes talk about a two-dimensional world by first talking about a three-dimensional world and then picking out a surface in this world. For example, we pick out a possible two-dimensional world in the following way by considering the surface of a sphere. Now, a sphere is a three-dimensional object—the collection of all points which are $n$ metres or less away from point $p_0$, the centre of the sphere. And the surface of the sphere is the collection of points precisely $n$ metres away from $p_0$. But of course, in a two-dimensional world which is the surface of a sphere, there is no point $p_0$ which is n metres away from every point! There is nothing mysterious or inconsistent about this process, though. The axioms of three-dimensional Euclidean geometry induce a set of properties and relations holding between the points on the surface of a sphere. By asserting these axioms, we entail that such a surface exists, and that the points on this surface instantiate certain properties, and that lines lying on the sphere which connect pairs of points have certain lengths. Now—take the surface of a sphere with all the properties and relations induced upon these points and lines, and subtract away all the other points of our Euclidean geometry. Since the shortest line between two points is no longer what it was, and since the distance between two points simply is the length of the shortest line, we have induced a new metric holding between the points in our world, and have thus arrived at a non-Euclidean world. Whether or not we have the means to describe this world intrinsically, without the need first to introduce a three-dimensional world, is irrelevant. We do successfully and unproblematically describe a particular non-Euclidean world by taking back some of the implications of what we earlier said.

Why indulge in such weasely behaviour if we can avoid it? Taking back things we have said before is often unhelpful and misleading, and is indeed somewhat weasely. Why not say exactly what we want to say first time round? Intrinsic descriptions of the surface of a sphere exist; and one can say that every *F* save Harry also *G*s. So why take back part of a picture which we earlier implied existed? Because sometimes we have to. Sometimes, we just cannot say what we want to say first time round. Sometimes, in order to communicate our picture of the world, we *have* to take back or modify part of what we said before. This is one of the things T and T* illustrate. There's nothing we can do in the language of T to say what we want to say about the sums. There *is* something in T* we can do in order to say something about the sums—but what we say commits us also to a load of extra platonistic entities, things which we do not believe in. Therefore, having asserted the sentences of T* and thus having captured what we believe about the sums, we make it clear that part of what T* says is false: so we go on to assert that there are no mathematical objects.

I am not defending any version of paraconsistency here. I do not think we ought to *believe* two inconsistent sentences. No—I am arguing for something much weaker: that sometimes, in order to present our (consistent) picture of the world, it is legitimate to take back details that were asserted earlier. Nor do I believe that even *this* is radical. It is common practice. Whilst almost all scientists will admit that they must quantify over numbers in order to formulate their scientific theories, almost all will go on to deny that there are such things as mathematical objects. Philosophers typically represent these scientists as engaging in double-think—denying by night what they believe by day. But it is surely uncharitable to regard so many scientists as hypocrites! Surely it is more charitable to think that we must have misinterpreted them.

But look at the kind of things they say: "The force between two massive objects is *proportional* to the *product* of the masses *divided* by the square of the distance"; "There is a one-to-one differentiable *function* from the points of space-time onto *quadruples of real numbers*"—how can we have misinterpreted them? By thinking that any theorist who presents a theory of the world must do so by asserting a set of sentences, each one believed by the theorist. This is our mistake. As soon as we allow theorists to take away details that were added before, to subtract parts of their earlier discourse, the theorists no longer appear to believe contradictory things. The mathematics is the necessary scaffolding upon which the bridge must be built. But once the bridge has been built, the scaffolding can be removed. It is surely more charitable to take scientists to be weasels rather than inconsistent hypocrites.

## 6 Putnam: a case study

Putnam is extremely unhappy with this way of proceeding. "It is like trying to maintain that God does not exist and angels do not exist while maintaining at the very same time that it is an objective fact that God has put an angel in charge of each star and the angels in charge of each of a pair of binary stars were always created at the same time! If talk of numbers and 'associations' between masses, etc. and numbers is 'theology' (in the pejorative sense), then the Law of Universal Gravitation is likewise theology." (Putnam 1975, p. 74)

I agree. It *is* hard to make sense of someone who first denies the existence of God and angels, and then goes on to say that God and the angels stand in some kind of relation to binary stars. But, even if we set to one side all the irrelevant and confusing details that are associated with talk of God and angels, Putnam's example is loaded.

My thesis is that, just as in telling a story about the world, we are allowed to add details that we omitted earlier in our narrative, so we should also be allowed to go on to take back details that we included earlier in our narrative. If Joe *first* asserts that God and angels do not exist, and *then* goes on to say something about the kinds of relations God and the angels stand to the stars, Joe not doing something which fits this model. We can't subtract some of the details of a picture if those details are not already there! But if we first assert that God has put angels in charge of the two stars *a* and *b*—and then alter that picture painted by denying that either God or the angels exist, then we *are* left with a picture of how the world might be: according to this picture, there are stars and there exist the binary stars *a* and *b*—but there are no Gods and angels. Of course, we might wonder why such a person didn't simply tell us that there were stars—and simply leave it at that. This is a fair accusation—but longwindedness falls far short of contradicting oneself, and certainly is no crime against rationality.

Now suppose that Joe tells the following story:

> In charge of each star is an angel, no two angels are in charge of the same star, and at the precise moment that each star is created the corresponding angel is also created. Moreover, the angels in charge of stars *a* and *b* were created at the very same time.
>
> Now I must modify something I said earlier—as a matter of fact, there are no angels, but apart from that, my story is correct.

Now, Joe never explicitly said that stars *a* and *b* were created at the very same time. Yet it is a consequence of his story. For, according to the first part of his story, angels are created at the very moment their stars are created. Since the angels in charge of *a* and *b* were created at the same time, stars a and b must also have been created at the same time. Now, by the second part of what was said, we must peel away the angels postulated in the first part. This leaves the stars, just as they were in the first part of the story, save there are no angels. And this leaves stars *a* and *b* still created at the same time.

In this case, there is no part of Joe's story which he could simply omit. Each sentence in the story must remain if his story is to have the implication that stars *a* and *b* were created at the same time. Of course, Joe is still being a little dim. For, though there is no *sentence* of his story which can be omitted, there is still a sentence he could assert which would save him the trouble of talking about God and angels; namely "the stars *a* and *b* were created at the same time". But again, Joe has committed no crime against rationality. Again, he's been longwinded and introduced fictitious entities he did not have to introduce. But for all this, it's quite clear what picture he has painted of the world, and again, it is not on to claim that Joe is being irrational.

Moreover, as the case of T and T* showed, there are cases where it is not even right to accuse Joe of being long-winded. There are cases when Joe just may not have the linguistic resources to eliminate his talk of abstract entities. At this point, quantifying over objects he doesn't believe exists *is* sensible—for it is the only means Joe has for saying what he wants to say. In these circumstances, Joe is forced to be a weasel.

One might think that, unless we could find a way of expressing Joe's beliefs not committing us to abstracta (for instance, by adopting a language which took the quantifier "there are finitely many *F*s" as primitive) then Joe's picture of the world is simply ineffable. But this is a mistake. By taking back some of the consequences of his earlier sentences, Joe succeeds perfectly well in communicating his picture of the world.[9]

[9]If T* + {there are no such things as sets} does not entail the existence of sets, what exactly *does* it entail? How should a logician interpret those theorists who take back part of what was earlier asserted?

It might look as though there is little for the logician *can* do. After all, although Joe believes that the sums are as T* says they are, there is no set of sentences of T* which correctly captures Joe's belief. Logicians often represent the content of a theory with the set of interpretations which make the theory true. But if our theory includes sentences which contradict each other then, on this model, the theory will be empty. We need a better model.

Consider a particular interpretation M of T*. Since M contains objects which fall under the predicate "*x* is a set", according to M, there are such things as sets. But now consider M^. M^ is got from M by deleting all the objects in M's domain which fall under the predicate "*x* is a set". M^ is a substructure of M. M^ says the same things about the nominalistic part of the world as M does, but unlike M, M^ "says" that there are no abstract objects. Let Δ be {M^ | M † T*}.Then if logicians want to model or study the content of T* + {there are no such things as sets}, then Δ is the correct candidate. Δ represents what a believer in T* believes and gives flesh to the idea that one can take back the implications of an earlier part of a story. The process of moving to M^ from M is the formal correlate of removing some of the details of what was earlier asserted.

## *7. Against the indispensability argument (2): the aesthetics of physics*

In this final section, I consider the second wing of the indispensability argument mentioned in §1. In this version, we are told that the kind of theoretical benefits postulating mathematical objects can bring are similar to the kind of theoretical benefits postulating concrete unobservables can bring. Although there may be no direct empirical evidence for either kind of entity we can see that postulating such entities results in an overall increase in the attractiveness of our scientific theories. If it is rational to accept concrete unobservables into our ontology for such reasons then it is equally rational to accept abstract ones. And, indeed, when we look at scientific practice we find that the postulation of such unobservables as quarks and space-time are justified by appealing to aesthetic considerations of simplicity, economy and elegance. Clearly, such arguments will only threaten those who accept some version of realism about the concrete unobservable objects postulated by physics—but given that I am such a realist, oughtn't I to stop weaseling and admit mathematical objects into my ontology for the same kinds of reasons I admit quarks and space-time?

I accept that the more attractive the theory the more reason we have, other things being equal, to prefer that theory to its rivals. Given a choice between two theories, both of which are compatible with the existing empirical evidence, we should measure the two theories' simplicity, their explanatory powers, their a priori probabilities, their economy and elegance etc., and place our belief in the theory that scores the most highly overall. Of course, it is a deep and difficult question how the various attributes that contribute towards a theory's attractiveness ought to be spelled out, and how these attributes are to be independently measured and weighed against each other. But there are a large number of cases where we can quite easily make such judgements and I accept that such judgements are rational. But whilst I agree that such principles have a role to play in theory choice, and that such principles may justify postulating quarks and space-time, it is a mistake to think that these principles justify our postulation of mathematical objects. To see this, let us consider a case where the postulation of mathematical objects apparently results in an increase of simplicity.

Many physical theories will need to describe the spatio-temporal relations that obtain between various physical objects. To do this, the theories will need predicates that can say that the distance between two objects is 3 metres, or $\frac{7}{11}$ metres, or $\sqrt{2}$ metres. Consider two theories that do this in two different ways: $T_1$ takes the two place predicates "*x* is-3-metres-from *y*", "*x* is-$\frac{7}{11}$-metres-from *y*", "*x* is-$\sqrt{2}$-metres-from *y*"[10]—and an infinite number of others—as primitive. Contrast $T_1$ with $T_2$: $T_2$ includes the single *three* place predicate "*x* is *r* metres from *y*", where the variable "*r*" ranges over real numbers, plus enough names for numbers to enable it to ascribe exactly the same distance relations as $T_1$.

At first sight, it may look as if we have to make a difficult decision between the two theories. $T_1$'s simple ontology is counterbalanced by the fact it must take an infinite number of predicates as primitive. This is in contrast with $T_2$ which has far fewer primitive predicates than $T_1$ but which pays for this by complicating its ontology. It seems as though considerations of simplicity fail to pick a clear winner: on a scale of simplicity, each theory has its own merits and demerits.

But this is a misapplication of simplicity. Although $T_2$ is able to *express* an infinite number of distance relations in a particularly simple way, there is no sense in which, according to $T_2$, the *world* is a simpler place than it is according to $T_1$. For, although $T_2$ uses fewer primitive predicates than $T_1$, there is no reason to think that $T_2$ actually postulates fewer fundamental relations than $T_1$. The fact that $T_2$ is capable of generating infinitely many distance predicates using merely a few primitives does not entail that the distance relations expressed by these predicates are themselves not primitive, irreducible relations. Indeed, in this particular case, although $T_2$ expresses the fact that a is $\frac{7}{11}$ metres from b by using a three place predicate relating a and b to the number $\frac{7}{11}$, nobody thinks that this fact holds *in virtue* of some three place relation connecting a, b and the number $\frac{7}{11}$. Rather, the various numbers are used merely to index different distance relations, each real number corresponding to a different distance. But if numbers merely index the relations then no conclusions can be drawn about the nature of the relations themselves—in particular, no conclusions can be drawn as to whether or not the distance relations are fundamental. Accordingly, for all $T_2$ says, it postulates no fewer fundamental distance relations than $T_1$ and so there is no reason to suppose that the kind of world postulated by $T_2$ is simpler than $T_1$.

I accept that considerations of simplicity play an important role in theory choice. But I prefer the hypothesis that makes *the world* a simpler place. For sure, all else being equal, I prefer the theory with the simpler ontology. For sure, all else being equal, I prefer the theory that postulates the least number of fundamental properties and relations. But the simplicity I value attaches to the kind of world postulated by the theory—not to the *formulation* of the theory itself. And while it's true that $T_2$ is capable of generating all the necessary distance *predicates* in a particularly simple way, I see no reason to

[10] The hyphens are there to emphasise the fact that the symbols "$\frac{7}{11}$","$\sqrt{2}$" appearing in the predicates are not to be taken as names for numbers.

think that *this* kind of simplicity should prompt us to prefer $T_2$ to $T_1$. Accordingly, I see no reason to think that $T_2$ is simpler in any relevant respect.[11]

The same is true for the theory T* discussed in section 3. Although T is capable of describing kinds of regions that lie beyond the descriptive powers of T*, nobody thinks that these extra kinds of regions exist *because* of the mathematical objects and properties postulated by T*. The mathematics is there to enable us to express possibilities that may be otherwise inexpressible but it plays no real role in simplifying our picture of the world.

Contrast the situation with typical concrete unobservables. Postulating quarks genuinely makes the *world* a simpler place. Under the quark hypothesis, various objects in the particle zoo do exist *in virtue of* the existence, properties and relations of quarks. By postulating few different fundamental kinds of objects and a few fundamental ways of arranging quarks, we can account for the existence of a wide range of apparently different kinds of objects simply, elegantly and economically. Moreover, this is a genuine, ontological account—the complex objects *owe their existence* to these fundamental objects and their modes of recombination. Accordingly, the principles do genuine physical work in simplifying our account of the world. Not so, the mathematical objects—or, at least, not so according to the standard mathematical platonist. It's wholly implausible to think that the sums that exist do so in virtue of standing in certain relations to abstract objects. Accordingly, though postulating the mathematical objects may simplify our theories, they do not simplify our account of the world. So, at least in these examples, the kind of pragmatic benefits afforded by postulating mathematical objects are quite unlike those we appeal to when we justify postulating concrete unobservables and their principles of composition.

Of course, there *may* be applications of mathematics that do result in a genuinely more attractive picture of the world—but defenders of this version of the indispensability argument have yet to show this. And certainly, the defenders need to do more than point to the fact that adding mathematics can make a theory more attractive: they have to show that their theories are more attractive in the right kind of way. That a theory can recursively generate a wide range of predicates, that a theory has a particularly elegant proof procedure, that a theory is capable of making a large number of fine distinctions are all ways in which mathematics can add to a theory's attractiveness. But none of these ways results in any kind of increase in simplicity, elegance or economy to our picture of the *world*. Until examples of applied mathematics are found that result in this kind

[11] If we were to follow philosophers of science such as Suppes (1967) and van Fraasen (1980) and identify theories with sets of *models* rather than sets of *sentences*, then there would be no danger of being seduced by simplicity of *formulation*.

of an increase of attractiveness, we realists about unobservable physical objects have been given no reason to believe in the existence of numbers, sets or functions.[12]

# *Appendix*

## *1. Proof that T has a model containing no infinitely coinfinite regions*

The important distinction between T and T* from the nominalist's point of view was that T* entailed the existence of a kind of region which T did not. Whereas T* guarantees the existence of regions which contain an infinite number of atoms and whose complement contains an infinite number of atoms, T does not. I shall call these special kinds of regions *infinitely coinfinite.*

Let M be a model having the following three properties: (1) M contains a countable number of atoms; (2) for every finite set of atoms S, M contains the sum of the atoms in S; (3) for every finite set of atoms S, M contains the sum of those atoms not in S; (4) M contains no other regions. Clearly, M has no infinitely coinfinite regions. I claim that M is a model for T.

Only the comprehension schema requires any work. So we need to show that, where $\phi(x, s_1 \ldots s_n)$ is a complex predicate and the $s_i$ are parameters, any region which is the sum of the $x$'s satisfying this predicate are in the model M.

Let f be a one-to-one function from the atoms onto the atoms. Given such an f defined upon the atoms we can extend this f in a natural way so that it acts upon the regions. If $a$, $b$, and $c$ are sent to f$a$, f$b$ and f$c$, then it is natural to say that f(the sum of $a$, $b$ and $c$) should be mapped onto the sum of f$a$, f$b$ and f$c$. More generally, f$(s)$ is the sum of those a such that a = f$x$, for some atom $x$ lying on $s$.

Notice that $f(a \cup b \cup c)$ can equal $(a \cup b \cup c)$ even when f is not the identity on $a$, $b$ and $c$.

Moreover, since $x < y$ if $fx < fy$, and $x = y$ if $fx = fy$. Since f is one-to-one and onto, f is an isomorphism. Accordingly, for any formula $\phi(x \ldots y)$, $\phi(x \ldots y)$ will be true if $\phi(fx \ldots fy)$ is true. In particular, notice that when $f(s_i)$ is the identity on the $s_i$, then $\phi(x, s_1 \ldots s_i)$ if $\phi(fx, s_1 \ldots s_i)$ for any $\phi$.

If we can show that, for any $a_1 \ldots a_n$ in M, the $x$ such that $\phi(x, a_1 \ldots a_n)$ sum to a region which is also in M, then we will have shown that the com-

[12] Versions of this paper were read out to the Moral Science Club in Cambridge and at York. I would also like to thank Otavio Bueno, John Divers, Rosanna Keefe, Dominic Gregory, John Harrison, Tor Lezemore, Adrian Moore, Christian Piller and Scott Shalkowski for helpful discussions and conversations.

prehension axioms are all true in M. Since each $a$ is in M, $a$ is either finite or cofinite. Partition the $a_1 \ldots a_n$ into a finite number of disjoint, non-overlapping regions, $b_1 \ldots b_k$ ($k = 2^n-1$ and $b_1 = a_1 \cap \ldots a_{n-1} \cap a_n$, $b_2 = a_1 \cap \ldots \cap a_{n-1} \cap a_n^{\perp}$, $b_3 = a_1 \cap \ldots a_{n-1}^{\perp} \cap a_n$, $b_4 = a_1 \cap \ldots \cap a_{n-1}^{\perp} \cap a_n^{\perp}$, $\ldots$ $b_k = a_1^{\perp} \cap \ldots \cap a_{n-1}^{\perp} \cap a_n$, where $a^{\perp}$ is the complement of $a$). Since the $a$'s are either finite or cofinite, each $b$ will also be either finite or cofinite.

We now return to the formula $\phi(x, a_1 \ldots a_n)$. Now, either there is an $x$ satisfying this formula which overlaps some cofinite $b$, or every $x$ which satisfies this formula overlaps no cofinite $b$. We prove that, in either case, the sum of the $x$ which satisfies $\phi(x, a_1 \ldots a_n)$ is either finite, or cofinite—and this establishes the desired result.

So consider the first possibility: suppose that there is some $x$ satisfying $\phi(f(x), a_1 \ldots a_n)$ which overlaps some cofinite $b$. Since $b$ is cofinite, it is infinite. Take any point of $x$ which overlaps $b$, and let f be a function which permutes this point with any other point of $b$, and is the identity everywhere else. Clearly, $\mathrm{f}(b) = b$, and so, since $b$ is a subset of each $a_i$, $\mathrm{f}(a_i) = a_i$. Now, as we saw above, $\phi(x, a_1 \ldots a_n)$ implies that $\phi(x, a_1 \ldots a_n)$. Accordingly, f$x$ satisfies our formula. This is true for any f we like, swapping a point of $x$ for some point of $b$. So, for any atom of $b$ we like, there is an f$x$ which contains that atom and which satisfies $\phi(x, a_1 \ldots a_n)$. So the sum of all the regions which satisfy $\phi(x, a_1 \ldots a_n)$ contains $b$. But $b$ is cofinite and any sum which contains $b$ must also be cofinite. So the sum of these $x$ is not an infinite coinfinite region—which is what we wanted to show.

Now consider the second possibility: suppose that every $x$ which satisfies $\phi(x, a_1 \ldots a_n)$ does not overlap any cofinite $b$.

So, every $x$ which satisfies $\phi(x, a_1 \ldots a_n)$ lies wholly within some finite $b$. So, every $x$ satisfying $\phi(x, a_1 \ldots a_n)$ lies within the union of these $b$'s. Since there are only a finite number of such $b$'s, and since each such $b$ is finite, the union of these $b$'s is finite. So every $x$ satisfying $\phi(x, a_1 \ldots a_n)$ lies within a finite region. So the sum of all these $x$'s lies within this finite region. So the sum of all these $x$'s must itself be finite. So, as before, the sum of these $x$ is not an infinite coinfinite region—which is what we wanted to show.

So, we know that, for any $a_1 \ldots a_n$ in M, the sum of the $x$ is such that $\phi(x, a_1 \ldots a_n)$ is also in M. And so the comprehension schema are indeed true in M. So M is a model for our theory.

*2. Proof that T is complete*

Kreisel and Krivine (1970) show that, by extending the language of the theory of atomic Boolean algebras, it is possible to eliminate the quantifiers from this theory. The following proof is nothing more than a modification of theirs. Atomic Boolean algebras are very like our atomic theory of mereology. The only essential difference is that atomic Boolean algebras contain a zero element, whilst there is no empty mereological sum. All the important philosophical points in the paper would go through if our base theory T were the theory of atomic Boolean algebras, and T* our mathematical extension. T is complete, it is unable to guarantee the existence of an infinite coinfinite region, T* is conservative over T, so T* entails no sentence in the vocabulary of T, yet it entails the existence of an infinite coinfinite region. However, it is philosophically nicer if our T is a nominalistic theory.

Introduce new predicates into our language, such as "$A_n y$" (meaning $y$ contains at least $n$ atoms) and add the obvious defining clauses for these predicates. By so doing, we end up with a new theory T^ as a result, but it is clear that T^ is a conservative extension of T. Now, Kriesel and Krivine's proof can be extended to show that T^ permits the elimination of quantifiers. Moreover, since T^ is nothing more than an extension of T by definition, the model M discussed above can be extended to a model of T^ simply by ensuring that the extensions of the new predicates satisfy their defining clauses.

Now, it is generally true that, if T^ permits the elimination of quantifiers and M is a model of T^, then T^ $\cup$ diag(M)[13] is a complete theory, whatever T^ and M may be. So, to obtain our desired result, we show (1) that if T^ $\cup$ diag(M) is complete, then T^ alone is complete and (2) that if T^ is complete, then T is complete.

T^ $\cup$ diag(M) is richer than T in two respects. Firstly, it contains new predicates of the form $A_n(x)$. Secondly, it contains new constants $c_i$—one for every region in the model M.

Suppose that T^ $\cup$ diag(M) proves S via proof P, where S is in the vocabulary of T^. Let $\Delta_1 \cup \Delta_2$ be the finite set of axioms of T^ and diag(M) used in the proof with $\Delta_1$ from T^ and $\Delta_2$ from diag(M). $\Delta_2$ is a finite set of atomic formulas true in M. It contains a finite number of new constants $c_j$ where these new constants are names of regions in M. Replace each constant $c_j$ with the variable $y_j$, where this variable appears nowhere else in $\Delta_2$. Conjoin the formulas so formed and add a sequence of existential quantifiers at the beginning (for instance, where $A_3(c_i)$, $c_i < c_k$ and $A_{17}(c_l)$ are the formulas of $\Delta_2$ which appear in the proof of S, $\exists w \exists x \exists y\, (A_3(w)\ \&\ w < x\ \&\ A_{17}(y))$ is the formula that results from this construction). It is clear that by following this construction on $\Delta_2$ we end up with a sentence in the language of T^. Call this sentence $\Sigma$. Suppose that $\Sigma$ is of the form $\exists x_1 \ldots \exists x_n(P_1\ \&\ \ldots\ \&\ P_m)$. By $n$ applications of existential instantiation followed by $m$ applications of conjunction elimination we

[13] Diag(M) is the diagram of M: add to your language a constant for every object existing in M—then diag(M) is the set of all atomic and negated atomic sentences of the expanded language true in M.

recover the sentences of $\Delta_2$. Since we are given $\Delta_1$ we can then simply mimic proof P and so derive S from T^ + {Σ}.

Now we show that Σ is provable from T^. M contains (i) a denumerable number of atoms, (ii) all regions in M are the sum of a finite number of atoms, and (iii) all regions which are the complement of a finite number of atoms—and nothing more. The existence of each such region is entailed by the theory T^. So every region which exists in M is proved to exist by T^. Accordingly, every model of T^ contains a submodel isomorphic to M. Now, since $\Delta_2$ is true in M, there are regions which satisfy $\Delta_2'$—the collection of formulas resulting from replacing the constants $c_j$ with variables) in M. So there are regions which satisfy $\Delta_2'$ in every model of T^. So there are regions which satisfy the conjunction of $\Delta_2'$ in every model of T^. So the existential closure of the conjunction of $\Delta_2'$ is true in every model of T^. But this is just Σ. So Σ is true in every model of T^. So, (by the completeness theorem for first-order logic) Σ is a theorem of T^.

This gives us the completeness of T^. So what about T? Well, T^ is merely an extension of T by definitions. It extends T only by adding new predicates and definitions for the behaviour of these predicates. It is known that extending a theory in this way results in a conservative theory. No new sentences in the old vocabulary are provable in T^ which were not provable in T alone. So, any sentence in the vocabulary of T which is provable in T^ is provable in T. So since T^ is complete, T is complete.

*3. Proof that every model of T* contains an infinite coinfinite region.*

T* contains instances of the comprehension schema for mereology including set-theoretic vocabulary. Moreover, T* entails that there is a one-to-one function f from the atoms into the ordinals. Now consider those atoms which are mapped into the even numbers by f. Since "*x* is an even ordinal" is expressible in T*, T* contains a predicate which is satisfied by some such set of atoms, and by the comprehension schema for regions, T* entails the existence of the mereological sum of such a set of atoms. Clearly, this mereological sum is infinite coinfinite.

*Department of Philosophy*
*Derwent College*
*University of York*
*York*
*YO10 5DD*
*UK*
*jwm9@york.ac.uk*

JOSEPH MELIA

## REFERENCES

Field, Hartry 1980: *Science without Numbers: A defence of nominalism.* Oxford, Basil Blackwell.

—1984: "Is Mathematical Knowledge just Logical Knowledge?", *Philosophical Review* 93, pp. 509–52.

—1985: "On Conservativeness and Incompleteness", in *Journal of Philosophy* 81, pp. 239–60.

Kreisel, G. and Krivine, J. 1970: *Elements of Mathematical Logic*, Amsterdam, North-Holland Publishing Company.

Morgenbesser, S (ed.) 1967: *Philosophy of Science Today*, New York: Basic Books.

Melia, Joseph 1998: "Field's Programme: Some Interference". *Analysis*, 58, pp. 63–71.

Putnam, Hilary 1971: *Philosophy of Logic*. New York: Harper

—1975: "What is Mathematical Truth?". In *Philosophical Papers*, vol 1. Cambridge: Cambridge University Press, pp. 60–78.

Suppes, P 1967: "What is a Scientific Theory?", in Morgenbesser (1967), 55–67.

van Fraasen, Bas 1980: *The Scientific Image*. Oxford: Clarendon Press.