Article
*The Withering Away of Formal Semantics?**

NEIL TENNANT

Why is semantics important? More precisely, why is *formal* semantics important? What is so special about the subject matter of formal semantics, and its methods, that sets it apart as a necessary component in any account of language?

By formal semantics I understand those broadly algebraic endeavours that set out to model the relationship between language and a (usually atomic) reality. The aim is usually to give a truth definition for a language provided with a formal grammar. Then, by appeal to the truth definition and the notion of truth-preservation, one characterizes a semantical relation of logical consequence. Depending on the logico-grammatical resources of the language, the models – those algebraic icons of reality – are more or less complicated.

At the simpler end of the spectrum lies the truth tabular semantics for classical propositional languages. There an interpretation is just a matter of settling a truth value for each of the atomic sentences (propositional variables). The recursion in the truth definition is then simply a matter of carrying out a truth table computation.

Nicely in the middle range are models of extensional first order languages with identity. Models there provide referents for names, functions for function signs, and extensions for predicates. Tarski's well known theory of satisfaction and of truth then provides clauses in an inductive definition of truth. The basis clause exploits the interpretative anchoring of the language in the model. Thus P(a) is true just in case the individual referred to by the name a lies in the extension assigned to the predicate P. More complex sentences and formulae are dealt with by recursive appeal to the characteristic truth-affecting contribution of the dominant connectives and quantifiers.

---

* I am grateful to both Martin Davies and David McCarty for comments occasioning revision of earlier drafts. Defects that remain are entirely due to me.

At the more complicated end of the spectrum are algebraic creations at the cutting edge of current research. We have the possible worlds structures of Montague semantics, reaching for constructions in fragments that other paradigms just can't reach. Here intensional operators such as the modal 'It is necessary that ...' or verbs of propositional attitude such as 'x believes that ...' are dealt with by exploiting the extra degree of freedom within such a structure. This is the freedom of shifting one's index or changing one's possible world, so as to be able to carry out evaluations at more than one index or world. The treatment aims to reduce the intensionality of these language fragments to the extensionality of a complicated ontology.

So much by way of introductory explanation of formal semantics. It is usually taken as but one half of the total project of characterizing a formal language. The other half is proof theory. Here one provides definitions of proof, by appeal to axiom schemata and rules of inference. Deducibility is defined in terms of the existence of a proof. Some proof systems are more natural than others; they usually achieve this by relying on rules of inference for each logical operator, isolating its unique deductive role in the language.

Once having given a semantics and a system of proof, the orthodox logician then proves *soundness* and *completeness* theorems. The proof system is sound if every sentence deducible from a set of premisses is a logical consequence thereof. It is complete if every logical consequence of a set of premisses is deducible therefrom.

Now my question at the outset was why formal semantics is needed. Despite the description of formal semantics just given, I am sceptical that any adequate answer is to be had. There are arguments to support that scepticism. They proceed as follows. Firstly, one presents as sympathetic an account as possible of the semanticist's received reasons for believing that formal semantics is both necessary and possible – necessary for our understanding of how language works, and possible with the methods currently employed. Then one shows the reasons to be vacuous. The project of formal semantics is not necessary, because all the insights and explanations within its compass are available within proof theory. And it is not possible with current methods, because they cannot mesh with reasonable requirements on any account of how language is learned. (Unfortunately, this final claim is one for which I shall not have space to argue in this paper. It is one that I have treated at book length in *Anti-Realism and Logic* (forthcoming from Oxford University Press, 1987).) My critique, if successful, will show how the division of labour between proof theory and semantics has resulted in no more than the parasitism of the latter on the former.

The General Triangle

When doing geometric proofs, one often draws diagrams as an aid to the imagination both in finding proofs and in following them. One draws, say, 'triangle ABC' on the page; one constructs lines parallel to its sides and passing through opposite vertices; one drops lines from vertices perpendicular to the other sides or the continuations thereof; one draws angle bisectors, or lines from the vertices to the midpoints of the other sides, and so on. It is now commonplace to observe that the diagram stands for no particular triangle; that it is only an heuristic to prompt certain trains of inference; that it is dispensable as a proof-theoretic device; indeed, that is has no proper place in the proof as such. For the proof is a syntactic object consisting only of sentences arranged in a finite and inspectable array. One is cautioned, and corrected, about an important misuse of that drawn triangle: the mistake of assuming as given information that is true only of the triangle that one has happened to draw, but which could well be false of other triangles that one equally well might have drawn in its stead.

'Triangle ABC' is no more than a placeholder in schematic reasoning. In the terminology of proof theory, it is a parameter in a proof that is to terminate with an application of the rule of universal introduction. The condition for the correct application of that rule is that the parameter should not appear in any of the assumptions on which the premiss of the rule depends. This can in turn be ensured by seeing to it that 'assumptions' about triangle ABC are covered by universally quantified statements about triangles. The covering will be by universal eliminations. Let me illustrate. The overall pattern will be this:

(1)

Let ABC be an arbitrary triangle — Every triangle is F

So, ABC is F — Every F is G

So, ABC is G (1)

Thus, if ABC is a triangle, ABC is G

Hence, every triangle is G

The first two steps are universal eliminations. The final step is universal introduction.

The logical behaviour of the parametric 'triangle ABC' in proof structures such as the example just given brings out explicitly what is behind our implicit awareness that we may not appeal to more information about the triangle than is available for any triangle whatsoever. Thus the 'general triangle' drawn on the page has no genuine role to play in the reasoning. Whatever is true about the correct use of such ploys can be recovered from a proof-theoretic account of various aspects of the inferential patterns among the relevant sentences.

The same holds for the use of Euler-Venn diagrams in monadic predicate logic. If one draws the appropriate diagram to help one appreciate the validity of the argument from 'all F's are G's' and 'there is an F' to 'there is a G', one may take oneself to be constructing the 'general model' in which the premisses are true, and checking that in this model the conclusion is indeed true also. But the general model is as much a chimera as the general triangle. It does no more than recapitulate information already available in the obvious natural deduction of the conclusion from the premisses.

I have mentioned these two examples of illusory 'semantics' or 'modelling' because my description of them is widely accepted, and can therefore provide a secure starting point for argument by analogy; and because the situation in semantics generally is analogous to that in the case of general models and general triangles. I intend to show that models have no peculiar and irreducible role to play in our understanding of language. The whole project of formal semantics rests on a confusion, writ large, like the one about Berkeleyan triangles. To the development of the general argument I now turn.

Syntax and Semantics

With the division of labour between syntax and semantics, it is always maintained, comes profit-sharing. The profits are soundness and completeness theorems. A soundness theorem shows that the proof system is correct, or in good working order. A completeness theorem shows that the proof system is strong enough to provide what it is called upon to provide: namely, a proof of any valid argument. The semantics, with its definitions of models and of (reference and) truth of (subsentential expressions and) sentences in a model, shows how language gets to grips with its subject matter. Then, by generalising over models in the definition of logical consequence ('Every model that makes all the premisses true makes the conclusion true also'), one captures the old idea that logical relations such as entailment or mutual inconsistency are a matter of form, not of content. Now syntax, or proof theory, comes in to do its blinkered work in obedience to this idea that logic is a matter of form, not of content. The rules of inference will be schematic on extra-logical vocabulary, and pay detailed regard only to the particular logical operators involved.

Moreover, the completeness theorem affords a very important *reduction* of the semantic notion of logical consequence to the syntactic notion of deducibility. It is not just a replacement via co-extensiveness, but a reduction. This is because logical consequence is defined with respect to *all* models. The quantification here is over all models of all cardinalities of domain. This is a very big 'all'. Only God would be able to survey all models of given premisses, in order to check whether they were models of the conclusion given in the argument to be appraised for logical validity. But proofs are finite and effectively checkable. And the completeness theorem tells us that if an argument is valid – that is, if its conclusion is indeed a logical consequence of its premisses – then there is a proof that will show this. By soundness, only one proof is needed; by completeness, there will be one. Epistemically, validity is now a one-off affair. The infinitary notion of validity of a given argument has been reduced to the existence of a suitable finite object – a proof of the argument.

This is obviously important. But equally obviously, neither the problem of epistemic access-via-reduction nor its solution would arise for one who did not grant the infinitary notion any place in his theorizing to start with.

Let us continue the orthodox account of why models and a completeness theorem are so important. The first strand was that validity is epistemically accessible – valid arguments have proofs. The second strand is that *unprovability*, likewise, is epistemically accessible. If an argument lacks a proof, then one who has a God-like point of view could run through all possible proofs (of which there are infinitely, albeit countably many), and see, unaided by inspection of objects outside the class of proofs, that this is indeed so. But human beings cannot run through all proofs in this way. So how is the fact of unprovability to be epistemically accessible? The answer makes appeal once again to the completeness theorem. There will be at least one *model* that makes all the premisses true, but the conclusion false. We have only to 'give', 'define', 'construct' or 'describe' such a model, and our job is done. The unprovable argument will have been *counterexemplified*.

So the epistemic method of *proof* is now complemented with the epistemic method of *counterexample*. Unprovability, too, is now a one-off affair. If an argument is unprovable, there is a model that will show this. By soundness, only one is needed; by completeness, there will be one.

Now I have no quibble with the stress on the importance of soundness. But it is a virtue that need not be laboured if one believes in the primacy of syntactic rules of inference in determining the meanings of the logical operators. Indeed, the consistent proof theorist will seek more refined and detailed accounts of soundness than ever dreamed of by his colleagues in the 'divided labour' camp. All provable arguments are going to be valid, by careful construction if not by outright definition. I shall not detain the reader with the details here. The point at this stage is only that we do not wish to take issue with the epistemic method of proof required in order to establish the validity of an argument. For the consistent proof theorist the validity of a bare argument (premisses followed by conclusion) is, one way or another, going to *consist in* the existence of a proof of a suitably refined kind.

My quibble is rather with the epistemic method of counterexample. *My contention is that counterexamples (and ultimately, all models) should follow the exit signs in the company of general triangles.* They have no claim on the best seats, and detract from the performance of proofs. I turn now to a description, and a proof-based re-working, of the method of counterexample in three areas:

(i) classical propositional logic
(ii) first order classical logic (finite models)
(iii) first order classical logic (infinite models)

The Revisionist Claim: Examples

I shall first illustrate my claim in the context of propositional logic – case (i) above. After the illustration I shall be in a position to formulate the claim in full generality. But the reader should beware that the illustration of the claim in case (i) cannot seriously be held as evidence for its assertion in general, because the illustration is degenerate. The truth of the claim in this case is not all that exciting or arresting, simply because this logic is decidable anyway, and there is a corresponding latitude of choice concerning what counts as 'giving proofs' and what counts as 'giving countermodels'. But the illustration is useful in fixing the ideas involved in the revisionist claim above, before it is formulated in full generality.

In case (i) a *model* is simply an assignment of truth values (T or F) to the propositional variables, which I shall take to be A, B, C ... etc. Suppose P(A,B,C...) is a sentence involving the variables indicated. Let t be a truth value assignment. By [P,t] I shall mean the *truth set of t on the variables occurring in P*. This is the set obtained as follows: A is in [P,t] if and only if t(A)=T, −A is in [P,t] if and only if t(A)=F; and nothing else can be in [P,t] except variables or their negations (where the variables concerned occur in P). The following theorem is easily proved by induction on the length of P, and is at the heart of Kalmar's proof of the completeness of propositional logic:

Theorem If t(P)=T then there is a proof of P from [P,t]; and if t(P)=F then there is a proof of absurdity (#) from [P,t],P.

Example: Let P be (A & B) v (C ⊃ D). Let t assign values as follows:

	A	B	C	D
t	T	F	F	T

It is easily seen that t(P) is T. For, since C is assigned F, the disjunct (C⊃D) receives the value T, whence the whole disjunction P comes out with value T. Now the truth set [P,t] which codes into sentences of the object language the information given by the truth value assignment t, is simply {A,–B,–C,D}. From this set of premisses – indeed, from the premiss –C alone – one can prove the conclusion P. (See Appendix.)

Consider now the sentence Q: (A v C) ⊃ (B & D). Under the assignment t above Q receives the value F. For A's truth makes the antecedent true, while B's falsity makes the consequent false. (So the argument P/Q is invalid.) Once again there is a proof tracing this reasoning. It has the undischarged assumptions A, –B (both drawn from the truth set [Q,t], which is the same as [P,t] above) and Q, and conclusion #. (See Appendix.)

I have chosen P, Q and t to make my illustration of the revisionist claim as easy as possible. The orthodox logician will exhibit, say, the assignment t as a countermodel to the (unprovable, hence) invalid argument P/Q. What he does 'looks semantic'. He is specifying semantic values for the simplest non-logical expressions, and applying clauses in a truth definition to arrive at the value T for the premiss P and the value F for the conclusion Q.

The revisionist logician sees matters in a different light. He gives a theory T – here, the truth set of t on the variables A, B, C and D – with the following properties:

(i) T is simply (the deductive closure of) {A,–B,–C,D}
(ii) T is obviously consistent (by inspection, or by its method of formation)
(iii) T is formulated in the same (object) language as P and Q
(iv) There is a proof of P from T (indeed, an intuitionistic one)
(v) There is a proof of absurdity (#) from T, Q (indeed, an intuitionistic one)

Trusting in the consistency of T, and in the soundness of proof, the revisionist logician concludes that there is no proof of Q from P.

Let us now abstract from the workings of this simple example. The general formulation of my revisionist claim, intended to undermine the special place that formal semantics has come to occupy, is as follows:

> In any logic held to be complete by those who divide labour between syntax and semantics, the method of giving a counterexample M to any invalid argument $P_1, \ldots, P_n/Q$ *can be replaced without loss* by the method, to be described below, of finding a theory T with the following properties:

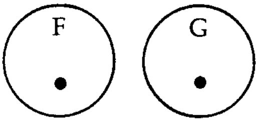
(i) T is effectively describable
(ii) T is obviously consistent
(iii) T is formulated in the same (object) language as are the premisses $P_1, \ldots, P_n$ and the conclusion Q
(iv) there is a proof of P_i from T, for each i
(v) there is a proof of absurdity from (T,Q)

The proofs mentioned in (iv) and (v) are available in the system that the orthodox logician would hold to be complete. Under the 'effectively describable' label would fall all first order axiomatisable theories – theories whose member sentences (theorems) can be effectively enumerated. But I am willing to admit a slightly more lenient specification than merely that the theory be axiomatisable. All I require is that the description of the theory T be at least as 'effective' as the description of the model M to which the orthodox logician is appealing in order to establish the invalidity of the argument $P_1, \ldots, P_n/Q$.

Let us now advance to an example of the revisionist claim at work with regard to finite counterexamples to invalid arguments of classical first order logic. This is case (ii) mentioned above. Again, we take as very simple example – the argument

(P_1) $\exists xFx$
(P_2) $\underline{\exists xGx}$
(Q) $\exists x(Fx\&Gx)$

The simplest counterexample to this argument would be the following two-element model:



The revisionist, however, sees in this model nothing more than an heuristic picture. What it really does duty for, according to him, is a theory, such as the one given by the following axioms:

$$Fa, \neg Ga, \neg Fb, Gb, \forall x(x=a \vee x=b)$$

This theory allows one to prove both premisses of the argument, and is also provably inconsistent with its conclusion. (For proofs, see Appendix.) Once again, these proofs are intuitionistic – as is the case in general. From a theory T categorically describing a finite model M, one can deduce every sentence true in M. (T categorically describes M just in case M is the only model of T up to isomorphism.)

Let us now summarize how the revisionist claim applies in the case of finite counterexamples to unprovable arguments in classical first order logic:

> Appeal to a finite model M can be replaced by appeal to a theory T with the following properties:
> (i) T is effectively describable (being the deductive closure of a finite set of axioms which are easily 'read off' from the given model M)
> (ii) T is obviously consistent (even without appeal to the model M as the source of the axioms of T; the consistency of T can be established by inspection)
> (iii), (iv), (v) as before

Against infinite models

So far my opposition might have been taking a tolerant line towards my demonstrations of the dispensability of countermodels when establishing the unprovability of an argument. We have, after all, been dealing with only the trivial fragments of propositional logic, and of that part of first order logic which, by virtue of having the finite models property, is decidable. It is to be expected, so the orthodox logician might say, that the more constructive, finitistic part of semantics can be displaced by surrogate methods from proof theory in the way I have been explaining. But, this line of defence continues, matters stand differently with full first order logic – especially when dealing with invalid arguments whose invalidity consists in there being an *infinite* countermodel, even though there be no finite ones. I shall counter this line of argument, and show

how the revisionist claim holds good even in the case of infinite models.

Let us take a well-known example: the ordering of the rationals. It is strict, transitive, linear, total, dense and unbounded. It is also non-trivial, in the sense that there are at least two elements to be ordered. Non-triviality, strictness, linearity and density suffice to make the domain infinite. These conditions can be expressed at first order by the following axioms respectively:

(P_1) $\exists x \exists y\, (x < y)$
(P_2) $\forall x \forall y\, (x < y \vee y < x \vee x = y)$
(P_3) $\forall x \forall y\, (x < y \supset y \nless x)$
(P_4) $\forall x \forall y\, (x < y \supset \exists z\, (x < z\ \&\ z < y))$

Take these as the four premisses of an argument whose conclusion is the claim that there is a greatest element in the ordering:

Q $\exists x \forall y\, (x = y \vee y < x)$

This is an unprovable argument, and the friend of formal semantics will want to provide a model in which the four premisses come out true, and the conclusion comes out false. Because the ordering of the rationals satisfies the four premisses but is unbounded both above and below, it would be a convenient countermodel to adduce for this invalid argument.

But in what sense has this countermodel been 'given'? Have we, so to speak, laid our hands on it any more directly than through the medium of such axioms as the four given above, and others (such as transitivity) that do not immediately feature in the argument? I would say not. Appeal to the existence of some model 'out there', against which the axioms can be laid to reveal their truth, is no better than taking a measuring rule from a hardware store and laying it against the standard meter rod in Sevres, in order to verify that the standard meter rod is indeed one meter long. Matters would be different if the intention were to verify that the measuring rule bought outside Sevres was itself one meter long. But so too with the case of our simple axioms and the model satisfying them. In my analogy, matters stand thus:[1]

---

[1] *Any* 'model' of the chosen axioms will do, from the point of view of the orthodox logician who wishes to establish the unprovability of the argument by appeal to a model. Interestingly, in the case we have chosen, model theory pronounces that there is only one countable model of the theory of the ordering of the rationals (one, that is, up to isomorphism). This is an old result due to Cantor. It follows from this that the theory in question is complete – it decides every sentence involving only '<' as an extralogical predicate. Moreover, being complete and axiomatisable, it is decidable. So it is a very well behaved theory. Not all theories with only infinite models are like this, however. As we know from the work of Gödel and Skolem, the set of all sentences 'true in the intended model of the natural numbers', is not merely undecidable, but also unaxiomatisable and *not even countably categorical* – that is, it has 'non-standard' models, in which there are elements other than the standard natural numbers 0, 1, 2, 3, ... but

simple axioms	$\leftrightarrow$	standard meter rod
the 'model'	$\leftrightarrow$	the measuring rule just bought
truth of the axioms in the model	$\leftrightarrow$	confirmation, by reference to the measuring rule, that the standard meter rod is one meter long

But let us return from this digression to the illustration of the revisionist claim in the case of infinite models. Our defence of the claim is going to look almost silly; but only at first sight. We had the argument $P_1, \ldots, P_4/Q$, where Q is the claim that there is a greatest element in the ordering, and the premisses $P_1, \ldots, P_4$ are the four axioms governing the ordering relation $<$ that we focussed on above. What is the theory T that does the trick for the revisionist? It is, dare I say, enough to take as its axioms $P_1, \ldots, P_4$ and Q', where Q' is so close to the negation of Q that the opposition may find my suggestions laughable:

$$Q': \forall x\, (\exists y\; y < x \;\&\; \exists y\; x < y)$$

On my suggestion, the proofs of the premisses P_i from T are trivial; and the proof that Q is inconsistent with T, given the form of Q', is almost trivial.

Why am I so confident that the revisionist can choose T in this way, without 'digging deeper' so as to have to do some proof-theoretic *work* to

---

whose presence does not disturb the truth of any of the sentences that are true when interpreted in the model consisting of *just* those standard natural numbers. What is the revisionist to make of all this? By eschewing talk about models, is he losing any important results and insights of foundational work this century? The answer will depend on how impressed one is by the supposedly substantive results of modern model theory, insofar as these are not recoverable proof-theoretically. The underlying point of view is that talk of models is mere handwaving, an aid to intuition, that momentarily, and for obscure psychological reasons, displaces talk of sentences and proofs involving them. It is remarkable how much of the model-theoretic story just given can actually be recovered in the latter realm of discourse. It is not as though the revisionist is deprived of the major foundational results of this century. The 'existence of non-standard models' is no more and no less than the consistency of number theory with an infinite set of axioms of the form $-@=\underline{n}$, for each numeral $\underline{n}$. Consistency follows from the very compactness of deducibility, in the familiar way. And as for Gödel's incompleteness theorem – it is constructively provable! Given any formal system for arithmetic, one can construct a sentence S such that there is no proof of S in the system, and there is no proof of –S either. No appeal to *models* is needed for this result.

Interestingly, the existence of *countable* non-standard models of arithmetic is a strictly classical result. In unpublished work David Charles McCarty has shown that *intuitionistic* arithmetic is countably categorical. This of course is from an intuitionistic metatheoretical standpoint, and derives from the intuitionistic unprovability of the downward Löwenheim-Skolem theorem. It is no accident that the considerations in this paper tend towards a proof-conditional, hence intuitionistic, 'semantics'. For a discussion of the downward Löwenheim-Skolem theorem, see our forthcoming joint paper 'Skolem's Paradox and Constructivism', *Journal of Philosophical Logic*.

get the premisses of the argument out of the theory T, and to get the conclusion of the argument contradicting T? The answer is simple: T as chosen already meets the desiderata set out above. It is effectively describable (did we not lay down finitely many simple axioms for it?). It is also obviously consistent. This, I maintain, is all that our 'grasp' of the 'structure' that we so readily picture to ourselves as

$$\cdots\overline{-2\quad -1\quad 0\quad 1\quad 2}\cdots$$

consists in. And it generates proofs as required, and described above.

We must not lose sight of the fact that picturing models to ourselves, or drawing diagrams, assures us of no better grasp of what is happening in the infinite case than does merely laying down that certain familiar axioms are to hold. We are so well versed in the use of these axioms, and so convinced by usual mathematical practice that they 'have models', that we do not pause to ask ourselves in what that supposedly external source for their 'satisfiability' *consists*. I maintain that it consists in no more and no less than our utter confidence that we shall never prove a contradiction from the axioms entertained.

I shall now press further with the argument to the effect that appeal to models can be supplanted by appeal to suitable theories. Some technical notions will come into play, such as categoricity and enumerability. The reader not acquainted with these will find them explained in my *Natural Logic* (1978). All the results that I shall invoke without proof in the arguments that follow are to be found either in that book or in Kleene's *Introduction to Metamathematics* (1962).

At this point the orthodox logician might reply along these lines:

> You have been clever in choosing the theory T that you did, in the last example. For it is *the complete and categorical* theory for the unique model that I had in mind when counterexemplifying that unprovable argument. But what about cases where I have a specific model in mind, but its theory is incomplete, unaxiomatisable and non-categorical? For example, what if I choose the standard model N of the natural numbers in order to counterexemplify an argument? How can you make me dispense with *that* kind of use of a model? What on earth would be the *theory* that could stand in its stead, along the lines suggested by your revisionist claim?

The answer is simple. To have been *that* specific about the model N, one would have had to demand the truth, at least, of all of Th(N). For otherwise, by Lindenbaum's Lemma, the sub-theory of Th(N) involved could be extended consistently by adopting any undecided sentence, or adopting its negation. This possibility immediately yields (by the orthodox logician's lights) competing, non-isomorphic models of the sub-theory for consideration. Yet this is just what the orthodox logician is supposed to be ruling out, by his concentration on *just* N. So we are supposing him to be concerned to establish the unprovability of some argument $P_1, P_2, \ldots$ / Q by appeal to the truth of all of Th(N). Th(N) is complete. Therefore, in order for the argument to be invalid, Th(N) would have to be inconsistent with Q and would have to contain P_1, P_2, ... Thus the orthodox logician's attempt at specificity regarding the choice of model N collapses into the revisionist move of merely taking Th(N) as somehow given, and then observing that it has the desiderata mentioned above (apart from describability). If the objection is raised that describability (i.e., axiomatisability) of the theory is important here, then the revisionist's retort is to point out that, in that case the 'argument' $P_1, P_2, \ldots$ /Q, involving, as it would now appear to do, all of Th(N) as its set of premisses, has hardly been 'given' anyway![2]

So the orthodox logician is frustrated in his attempt to have a unique model do work that appeal to proofs cannot do. If an argument has been *given* – at the very least, effectively presented – for which infinite models are *needed* in order to establish invalidity – then we could be dealing with a class of more than one model as counterexamples. For the revisionist logician who wishes to replace talk of models with talk of theories, the question now must be: is this class an elementary class? That is, is it the class Mod(T), for some theory T formulated in the object language? Or, more finely to the point: does it include an elementary sub-class? The answer, trivially, is affirmative. Take the argument $P_1, \ldots$ /Q. A counterexample to it is a model of $P_1, \ldots; -Q$. Taking $P_1, \ldots; -Q$ as the (axioms of a) theory T, we have that the class of models serving as counterexamples is Mod(T).

But this easy a defence of the revisionist claim is far too quick. It does nothing to solve the problem of how one can come to know on good grounds that the argument is invalid. That the theory T is consistent is no more obvious than that T has a model. So there is no epistemic advance on the matter of invalidity. So the dialectic of our position must be to ask the orthodox opposition to present persuasive instances where the description or definition of a model has led to epistemic advance; that is, where it has become *more obvious* than it was before that the argument in question is invalid. Or better, since it would not be good enough to deal with the opposition on adduced instances in turn, we should ask him to specify the general features of any situation which make it one of such epistemic advance; and we must then show that, in such a situation as generally described, it is *theories* and *proofs* that lie at the heart of the supposedly semantic contribution in this epistemic advance.

This places me in a somewhat difficult position. For in order to conduct the dialectic further, I have to provide the sort of general account asked

---

[2] Note that it is a further objection to the orthodox logician's strategem here that the complete theory Th(N) is not even countably categorical anyway, by his own lights. But the rejoinder worked out above does not even have to appeal to this fact.

of the opposition; and I can see no way to do this. Or, rather, the best I can do is follow a line of thought that originates in the best model existence theorems at the disposal of the orthodox logician, and see where that leads us.

The best such theorem asserts that every countable theory, if consistent, has a countable model. Due to Löwenheim and Skolem, a more refined version of the theorem states that for every countable theory T, and every model M of T, there is a countable *submodel* M′ of M satisfying T. Another refinement, due to Hilbert and Bernays, states that every axiomatisable consistent theory has a model *based on the natural numbers* in a particularly interesting way: namely, that every primitive predicate occurring in sentences of the theory can be assigned an extension in the natural numbers that is defined by a formula in the language of arithmetic that is equivalent both to a formula of the form $\forall x \exists y F(x,y,z)$ and to one of the form $\exists x \forall y G(x,y,z)$, where F and G are primitive recursive. More briefly: every consistent axiomatisable theory is satisfiable in the natural numbers upon Δ_2-interpretation of primitives. The interest of this result lies in the fact that these interpretations are relatively low down in the hierarchy of arithmetical complexity. The extensions are as close to being effectively enumerable as they can be.

It may be thought that this last version of the model existence theorem puts the orthodox logician in a reasonably strong position. Here is a quite narrowly restricted class of models in which he can successfully search for a model of $P_1, \ldots P_n, -Q$ in order to counterexemplify the argument $P_1, \ldots, P_n/Q$.

But my question would be: how would he know when he had such a model? Does one simply take the Δ_2-formulae on offer as the interpretations of the primitives involved, and 'see' that the premisses $P_1, \ldots, P_n$ thereby come out true, but the conclusion Q comes out false? Of course not. To make use of such a model the orthodox logician will have to substitute his Δ_2-formulae for the primitives, to get, say $P_1', \ldots, P_n', Q'$; and then he will have to *prove* each P'_1 from Peano's postulates (or whatever extension of those postulates he and his opposition find acceptable), and he will likewise have to prove that Q′ is inconsistent with Peano's postulates. In other words, with Peano's postulates playing the role of the theory T in my revisionist claim, he will have to do exactly what the claim says is involved in showing that there is no proof of Q from $P_1, \ldots, P_n$.

Now it is interesting to note that, if limited to any particular axiomatisable consistent extension of Peano arithmetic as one's 'base theory' T, this method cannot in general be assured of success. This is because if it were, first order logic would be decidable, contrary to Church's theorem. To see this, consider simply the problem of deciding theoremhood, or logical truth. We already have in any standard axiomatisation of first order logic a method for effectively enumerating the logical truths. (This is simply what the completeness theorem tells us). Now *if* we also had a method for enumerating *falsifiable* formulae (or, equivalently, satisfiable formulae of the form –P), *then* we would have a method for deciding logical truth. It would consist simply in setting the two enumerations going, and waiting to see in which list the sentence in question occurred. It has to occur in one or the other, and do so within a finite time. Whence the logical truths would form a decidable set. So now we must enquire after the possibility of enumerating satisfiable formulae of the form –P. We would get such an enumeration simply by combing for the negations occurring in a more general enumeration of all satisfiable formulae. But the Hilbert–Bernays result tells us that any satisfiable formula P will be true in the natural numbers upon some $\triangle_2$-interpretation of its primitives. What we do then is enumerate all $\triangle_2$-formulae. (This is possible, given completeness of the logic and the decidability of the syntactic forms concerned.) By well-known methods, we try out every possible substitution of $\triangle_2$-formulae for the primitives in P. As the list of $\triangle_2$-formulae grows longer, the possible substitutions grow more numerous; but every substitution will eventually be tried out. Now, *if* the truth of P upon any such substitution (so I should say: the truth of the relevant P′) were always within the compass of an axiomatisable theory T, this could be effectively discovered. So we would always be assured of reaching some substitution of $\triangle_2$-formulae (already established as $\triangle_2$) into P to get P′, and a proof of P′ from the base theory T. Hence there would be precisely the sort of complementary enumeration of falsifiable formulae, which, along with the enumeration of logical truths already given by a standard axiomatisation of logic, would render first order logical truth decidable – contrary to Church's theorem. Thus the assumption in our reasoning that is to be rejected is the assumption that the truth of P′ lies within the compass of proof from T. In brief: the Hilbert–Bernays method of giving models does not exactly *give* them! There must be at least some consistent P whose satisfiability in the natural numbers upon $\triangle_2$-substitution for primitives lies beyond the reach of Peano arithmetic. And in general, for any consistent axiomatisable extension T of Peano Arithmetic, there will be such formulae P.

But let not the orthodox logician think that the revisionist's claim is hereby undermined. For the situation is as bad for the former as it would allegedly be for the latter. Whatever one's theory T of arithmetic, as based on current axioms, such recalcitrant P in question will be just as recalcitrant for the orthodox logician as it will be for the revisionist. Their epistemic predicament regarding the invalidity of an argument $P_1 \ldots P_n/Q$ for which $P_1 \& \ldots \& P_n \& {-Q}$ is thus recalcitrant, will be *the same*. *But* – and this is the important point – whenever the orthodox logician *does* claim success (with good reason) in his search for a countermodel, the revisionist can *match* this by entirely proof-based methods: namely, a substitution of suitable $\triangle_2$-formulae for the primitives involved, and proofs of each P'_i from his theory of arithmetic, as well as a proof that Q′ is inconsistent with that theory.

I can think of only one remaining argument, or intuition, that has to be met from the orthodox opposition. This is the appeal made to one's 'intuitions about the standard model' in the course of the so-called 'semantic argument' for the truth of the undecidable Gödel sentence of arithmetic. The situation is familiar: for a given axiomatisable, consistent theory T of arithmetic, one constructs an undecidable Gödel sentence of the form $\forall nG(n)$. One is able to prove each instance; G(0), G(1), G(2), ... ; but one cannot prove $\forall nG(n)$; nor can one refute it. The 'semantic argument' consists in this: since each instance is provable, it is true; whence, since the instances deal with all the standard numbers, the universal quantification $\forall nG(n)$ is true in the standard model.

But there is no especially *semantic* insight involved here. One is only applying mathematical induction one level up linguistically. To be sure, if one artificially delimits in advance what is to count as a formal proof of arithmetic, then assertability is going to start bursting at the syntactic seams. But if one leaves proof as open-textured, precisely this sort of example will lead one to see that the need may arise for an extension of methods from, say, Peano arithmetic. But there is a clear sense in which it is still *induction* that is involved, albeit at a higher level. So once again, I conclude that there is nothing especially *semantic* going on in the argument for the truth of the Gödel sentence.

So far my discussion has been confined to first order logic, which is complete. But what of *incomplete* systems, such as second order logic? It is often argued that logical incompleteness secures a special emphasis on the semantics of the language involved; that one may account for its workings, so to speak, only by concentrating on the semantics, to the total exclusion of any proof theory. This is not an uncharitable characterization of work in Montague semantics, for example; and the apologetic just sketched is explicit in an introductory text such as D. Dowty et al, *Introduction to Montague Semantics*, 1981.

I shall be brief in my rejection of incompleteness as good reason to pay any attention at all to model theory at second order, let alone reason to ignore entirely what proof-theoretic techniques have to offer in the way of analysis of meaning. The assertion of incompleteness rests crucially on merely claiming to have a definite conception of the totality of natural numbers. How is this? There are two ways of proving incompleteness of second order logic. The first establishes merely that second order logical consequence is not compact; whence there can be no logic complete on arbitrary sets of premisses, since proofs involve at most finitely many premisses. The second establishes that even second order logical truth is not axiomatisable.

Let us look at each of these proofs in turn. For the first, take the second order theory T of N, the standard model of the natural numbers; and, using the new name @, say, form the infinite set of inequalities of the form $-@ = \underline{n}$, for each numeral $\underline{n}$. It is clear that each finite subset of the union of T with this set of inequalities has a model. Simply take, as the model, N supplemented with the stipulation that the name @ is to denote the first number whose numeral does not occur in any of the (finitely many) inequalities in the finite subset chosen. Yet – and here is where the intuition of the naturals comes in crucially – it is 'equally obvious' that T *along with* the full set of inequalities has no model. This is because second order T, with the induction axiom, determines N as its only model (up to isomorphism); and there cannot be any denotation, within the model N, for the name @, since this is ruled out by the set of inequalities involving that name and all the available numerals. Thus the second order argument

$$T, -@=0, -@=\underline{1}, -@=\underline{2}, \ldots / \#$$

is valid. But any argument obtained by dropping all but finitely many of the premisses is *in*valid. Hence second order logical consequence is not compact. Let me emphasize once more how the crucial step is the one where we say that the second order theory T has only N as a model. More specifically, we take the second order induction axiom to rule out any model with elements not denoted by a numeral. In the axiom

$$\forall P\,(P(0) \supset (\forall x\,(P(x) \supset P\,(s(x))) \supset \forall y P\,(y)))$$

we take as a 'substitution instance' for P (x) the property:

x is denoted by numeral $\underline{n}$, for some natural number n

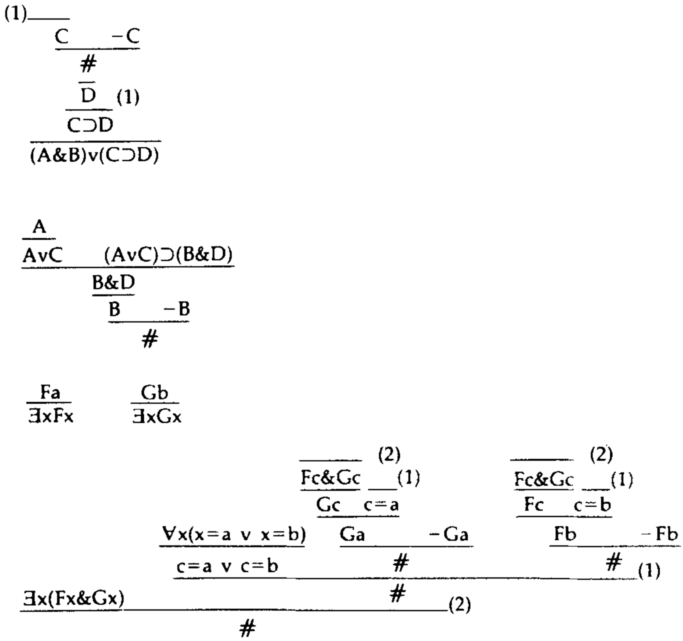
Exactly the same substitution in the second order induction axiom is at work in the second proof of the incompleteness of second order logic – the one that shows that even second order logical truth is unaxiomatisable. Assume we already have the proof that the first order theory Th(N) is unaxiomatisable. Let T now be the conjunction of the finitely many axioms of second order arithmetic. Once again, we observe that T has only N as a model. Hence those first order sentences that are second order logical consequences of T form exactly first order Th(N), which Gödel showed to be unaxiomatisable. Since there is an effective method for telling whether a sentence is first order, the second order logical consequences of T must form an unaxiomatisable set. Moreover any effective enumeration of all second order logical truths would, by checking for enumerated sentences of the form (T⊃S), yield an effective enumeration of all second order logical consequences of T, which as we have seen is impossible. Thus second order logical truth is unaxiomatisable.

The whole burden of 'semantic insight' in these two proofs of the incompleteness of second order logic is borne by the claim that the second order induction axiom, via the substitution instance above, ensures that only the standard natural numbers are in the domain. But, as remarked earlier, this is perfectly comprehensible and representable from a proof-theoretic point of view. One is simply *applying induction*. That is a move governed by certain syntactic constraints. There is no mysterious transcendent semantic intuition at work here, delivering itself of results beyond the reach of formulable rules of proof.

I end therefore by concluding that model theory is parasitic upon proof theory; that it is not a special repository of otherwise inaccessible insights;

and indeed that by concentrating wholly on the infinite algebras contemplated in later developments such as Montague semantics, one will be treading a path whose initial segment, as we have just seen, is a well worn rut that is wearing ever deeper, but getting nowhere. We must learn instead to navigate by the very features of the land itself. I believe that careful attention to the inferential moves we make not only within English (thought of for a moment as the object language) but also within English-plus-ZF *when doing Montague grammar* (testing both our own logical intuitions and the adequacy of semantic analyses within that tradition) will reveal all that we need in order to specify and systematize meanings. Perhaps this last claim is too much to base on the negative conclusions justified by the discussion above. But the reader will appreciate that limitations of space prevent me from making this the occasion to develop a positive case for the untapped resources of a proof theoretic approach to natural logic. I hope, however, to have softened up traditional resistance to the line I would wish to see followed.

Appendix



Department of Philosophy
The Faculties
Australian National University
Canberra ACT 2601

MS: accepted May 1986