Logique & Analyse **139-140** (1992), 239-249

LOGICAL CONSEQUENCE AND MODEL-THEORETIC CONSEQUENCE

Greg O'HAIR*

1. *Introduction.*

Since Tarski's work on the notions of truth and consequence in the 1930s, model theory has flourished alongside the older tradition of proof theory. It has led to many impressive technical achievements, and a new, semantic perspective on logic. Moreover, the model-theoretic account of logical consequence based on Tarski's work has been taken to provide a successful analysis of the common-sense or pre-theoretic notion. Unsurprisingly, the influence of Tarski's work has spread farther afield. Many people working in areas such as A.I., Cognitive Psychology, Epistemic Logic, Epistemology or Semantics of Natural Languages have adopted the model-theoretic concepts as part of their intellectual equipment for investigations of language, thought or reasoning.

In his recent book, *The Concept of Logical Consequence* [Etchemendy 1990], John Etchemendy has subjected this account to a powerful critique, exposing confusions and bad arguments that have contributed to the illusion that it captures the everyday notion.(1) He shows that it does not yield an extensionally adequate account of logical truth or consequence for arbitrary languages, thus forfeiting also the weaker claim to give a new, technically precise concept coinciding (roughly) in extent with the ordinary notion.(2) What about our strong feeling that, for some (artificial) languages, the account is *guaranteed* to be correct? Etchemendy shows in chapter 11 that though the guarantee does indeed hold in special cases, it does so in virtue

* This paper was written while I was on leave at the Center for the Study of Language and Information, Stanford University. I have benefitted from discussion with John Etchemendy, Jon Barwise, Chris Mortensen, Ed Coleman and Graham Priest, and from a seminars I gave on [Etchemendy 1990] at Stanford and at the University of Adelaide.

(1) See, for example, his discussion of the confusion between "representational" and "interpretational" conceptions of semantics in chapter 2 (and later), and his discussion of "Tarski's Fallacy" in chapter 6.

(2) The bulk of the relevant discussion here occurs in chapters 7 through 9.

of there being a proof system that is *intuitively* sound!

In this paper, I make no attempt to evaluate Etchemendy's critique. Rather, my aim here is to explore its implications. (I hope the issues will be of interest even to those who do not accept the critique.) First, I raise some issues for those who wish to continue to use the familiar methods where appropriate. Secondly, I consider an alternative conception of semantics going back (perhaps) to Wittgenstein, which Etchemendy suggests has been confused with the standard model-theoretic approach. Next I consider a related issue, that of our continuing dependence (if Etchemendy is right) on the intuitive notion of logical consequence. This points to the need for investigation of the intuitive notion. I conclude with a look at some issues for such a study —for example, the relation between logical consequence and modal notions such as that of necessity, and the apparent failure of the model-theoretic approach to *reduce* the notions of logical truth and consequence to non-modal notions.

2. *Using the model-theoretic concepts.*

Of course, the *technical* achievements of model theory survive Etchemendy's critique intact. They were never part of his target, and indeed, unlike some early critics of Tarski, he is in sympathy with the aim of using mathematical methods in studying these matters. However, his analysis does raise some serious questions about the *application* of the model-theoretic methods. Remember that Etchemendy has shown (assuming, for our purposes, the correctness of his account) that, in the general case, the model-theoretic account can be expected to give extensionally incorrect answers, perhaps substantially and systematically so. That is, the account may:

- *undergenerate*— fail to declare something to be a logical consequence (of other given items) when it is so, or
- *overgenerate*— declare something to be a logical consequence when it is not so.

On the other hand, Etchemendy has shown how, under certain conditions the model-theoretic account may be used without fear of these errors. What are these conditions? Essentially, there are two, and each deserves some attention.

2.1 *A new role for completeness theorems.*

Modifying and building on an earlier idea of Kreisel's, Etchemendy is able to show that, where we have a proof system that is *intuitively* sound and in addition, according to the model-theoretic semantics, complete, we may be assured that the semantics will deliver answers that are intuitively correct. There are two novel twists here. One concerns the kind of logical consequence involved, and will be discussed below. The other concerns the role of the completeness theorem, in enabling an intuitively correct proof system to guarantee that the (model-theoretic) *semantics* is correct! We are so used to assuming that soundness and completeness theorems "justify" the proof system rather than *vice versa* that this can come as a shock!

After absorbing the shock, we are left with the problem that, notoriously, a completeness theorem is not always available. Sometimes, we just have to live with that fact, a theme to be explored below. However, Jon Barwise makes the useful observation that we should not give up too soon: there are times when some modifications may result in a completeness theorem becoming available after all [Barwise 1991]. The idea is as follows. By hypothesis, our proof system is incomplete. Traditionally, this would have been seen as a defect of the proof system, or as a problem for provability with respect to some kind of language. Thanks to Etchemendy, we now see that the model-theoretic semantics may be at fault: if some "valid" sentences are not provable, it may be because they are not really (i.e., intuitively) valid after all! We took them to be (model-theoretically) valid because they are true in all models; and,indeed, so they are, *but only relative to our model-theoretic set-up*. We may alter this set-up in either of two ways: by keeping the same collection of models, and changing the semantic arrangements (an idea suggested by Peter Aczel), or by moving to a larger collection of models, including some where the troublesome sentences are false. Provided that we can do this, our original proof system is now complete (anything still valid after the changes to the model theory was already provable), and Etchemendy's guarantee applies after all.

Another strategy, of course, is to try to strengthen the proof system, so that sentences hitherto unprovable can now be proved. Here, too, we can (in principle) do more than might have been realized. For, as long as our extensions to the proof system remain intuitively sound, they need no further virtues to the job we need them for —guaranteeing the correctness of the model-theoretic semantics. In particular, they do not need to be 'cognitively natural', tractable or finitary.

Thus, we have new motivation for seeking, and hopes of (sometimes) finding completeness theorems where previously we thought there were none to be had. This is not a universal panacea, however, and we need to consider what to do in the face of continuing incompleteness. Before dealing with this, though, we need to look at another condition involved in Etchemendy's guarantee.

2.2 *Two kinds of consequence.*

Consider the significance of completeness theorems again for a moment. On the traditional conception, they showed that the proof systems in question were powerful enough to encompass all the valid sentences or consequences. On the new perspective, they can be used to show that the model theory is not too 'powerful', that it does not overgenerate. What about the converse problem, that of undergeneration? The methods considered so far do not seem to allow of a general solution to this problem. However, in one of the high points of the book, Etchemendy shows how, in the spirit of modern algebra, the problem can, for some purposes, be defined away. The idea is to think about the logical properties, not of one interpreted language, but of a whole *family* of such languages, the family covered by the model theory. Take the logical truths (or consequences) common to all the languages in the family. Etchemendy shows that, assuming as before that we have a completeness theorem for an intuitively sound proof system, a Kreisel-style argument will guarantee that the model theory delivers exactly the right answers about this collection.

This result is, I believe, an interesting and important one. It points to a useful distinction, between what we might call 'algebraic' consequence and 'single language' or 'single structure' consequence. With the algebraic tendency in modern mathematics, much work has been concerned essentially with whole families of structures, for example, all those that satisfy the axioms for groups or rings. The same tendency is found, of course, in modern topology, where whole families of spaces with certain properties in common are investigated at once. It seems illuminating to think of the algebraist as exploring the logical (as well as mathematical) properties of such families —the (algebraic) consequence relation for the languages (or structures) of, say, group theory.(3) The down side is that sometimes, as

(3) This may shed some light on model theory's particularly close connection with algebra, notwithstanding its many diverse applications to other areas.

in set theory, number theory or (one conception of) Euclidean geometry, we are really concerned with one structure, or one interpreted language used to talk about that structure. Though these familiar examples are from mathematics, this was largely to balance those of algebra and topology, and to remind us that there seem to be both kinds of study within mathematics itself. There must be many examples from the cognitive sciences, from Linguistics or A.I., say, where essentially one structure, or one interpreted language, is in question at a given time. This is an appropriate moment at which to turn to the alternative approach to semantics sketched by Etchemendy, since it offers help at this point.

3. *Representational semantics.*

On the standard, interpretational view, we think of a sentence, say, 'Snow is white' as being true in one model and false in another in roughly this sense: a sentence having in itself no particular meaning (an uninterpreted string) is given one interpretation under which it is true, and another under which it is false. So the 'world' (some domain of 'objects', and functions and relations over them) is held fixed, while the meaning varies from model to model. As Etchemendy notes, we could think about models differently, representationally: one model could represent a 'world' in which snow is indeed white (and hence our sentence is true), and another represent a different world, in which snow is, say, red (and our sentence, still having its customary meaning, is false). So here the meaning is held fixed, while the 'world' varies. Perhaps Wittgenstein had this conception in mind. At any rate, there are these two possible dimensions of variation.

The two approaches have a very different 'flavour', but can we be sure that they are not really equivalent, two perspectives on the same model-theoretic reality? Etchemendy shows that, while in special cases they deliver equivalent semantics (declare the same items to be logical truths or valid arguments), in general they diverge. So we have two genuinely different ways of doing semantics. Which should we prefer? And how would representational semantics work? More work needs to be done on such questions.([4]) In terms of our discussion in the preceding section, one point stands out: the alternative, representational approach offers solid advantages

([4]) Barwise and Etchemendy made a start on what Barwise describes as a 'rational reconstruction' of semantics along representational lines in [Barwise and Etchemendy 1990].

where the conditions for Etchemendy's guarantee do not apply. The alternative, representational approach looks promising here, since it does not require a completeness theorem and makes sense in relation to single interpreted languages. So, like the intuitive notion, it is available when the interpretational approach isn't; while, like the interpretational approach, it offers the advantage over the unaided intuitive notion of systematic, mathematical methods. True, it is relatively unexplored, and we need to understand it better. Still, it begins to look as though we can have the best of both worlds! A more sobering perspective will emerge in the next section.

4. *Dependence on the intuitive notion.*

Reflecting on the discussion above, we can see two ways in which, if Etchemendy's analysis is correct, we still depend upon the intuitive notion of logical consequence. The first concerns its role in, so to speak, underpinning the model-theoretic semantics, however satisfactory and successful the latter might be. The second concerns certain limitations on the availability or appropriateness of the model-theoretic notions in some cases.

4.1 *Underpinning the model-theoretic semantics.*

In the case of interpretational semantics, this role is a striking implication of Etchemendy's account, emphasized already above. As we saw in looking at the 'new role' for completeness theorems, we need the intuitive conception to warrant the (intuitive) soundness of the proof system that (via the completeness theorem) warrants the correctness of the interpretational semantics.

It should be emphasized that representational semantics, too, needs the intuitive notion. This applies, indeed, in the very setting up of a structure of models.

We might try to distinguish between the role in the two cases, in something like the following way. In the former case, the intuitive notion turns out to be needed to provide an 'external' guarantee for a significantly different concept. In the latter case, the role is more intimate, even 'internal' to the very idea of representational semantics. Be that as it may, in both cases, the role is an ongoing one: however appropriate and successful any application of either variety of model-theoretic semantics may be, the intuitive notion and its role are there in the background. This raises 'foundational' questions, for example about our 'justification' for both the intuitive notion and judgments based on it.

4.2 *Limited availability of model-theoretic semantics.*

Another source of dependence on the intuitive notion lies in the limited availability or appropriateness of the model-theoretic concepts. What is involved here is more than a temporary failure to develop or adapt such concepts in ways that make them suitable for specific areas, as in the case of various modal or intensional logics. Rather, there seem to be certain limitations of principle involved. Once again, this seems to apply to both interpretational and representational semantics, though in different ways in the two cases.

In the case of interpretational semantics, one such limitation stems from the need for a completeness theorem in order to validate the model theoretic set-up. Another concerns the distinction drawn earlier between algebraic and single-structure consequence. This suggests that where our concern is the study of a single structure, interpretational semantics will never quite fit the bill. This in no way detracts from the usefulness of model-theoretic methods in teaching us more about the properties of a language or structure. But the limitation is there. How significant we think it is will depend, for example, on whether we think that important parts of mathematics are best understood as the study of single structures.

That, of course, is where representational semantics seems to step into the gap exposed by Etchemendy's analysis. However, things are not so simple. It seems as though there will be times when neither interpretational nor representational semantics will be fully appropriate for our purposes. The classic case of elementary number theory provides an illustration. This case is, famously, a problem for the standard approach. It fails to meet either condition for Etchemendy's guarantee, since it is incomplete, and since we are really concerned with the one interpreted language, the one structure of natural numbers. However, in this case, we cannot turn to representational semantics for help. The reason is that we cannot make sense of the idea of alternative models, representing 'possible worlds' where, say, $2+2 \neq 4$. As Etchemendy says, "That way madness lies"(62).([5]) The problem in

([5]) There may be some irony in the fact that this case continues to be a source of problems for model-theoretic semantics so many years after Tarski used it to good effect against earlier approaches.

this case is that arithmetic consists of necessary truths. Obviously, there are other examples. Perhaps further study may reveal different cases, where for one reason or another representational semantics is not appropriate.

4.3 *Proof theory to the rescue? The case of number theory.*

If Etchemendy's main points are granted, it may be conceded that no formal theory of arithmetic quite hits the mark. In fact it might be said that this was well-known already. In the case of first-order theories, there is the existence of non-standard models, as well as the fact that any particular such theory is both too weak (by Gödel's first incompleteness theorem) and too strong, in the sense of having a lot of unused power and complexity by comparison with informal number theory. Second-order arithmetic is interestingly different, but brings its own problems...

Still, it might be claimed, none of this means that we are thrown back on some hazy and questionable, intuitive notion of 'arithmetic consequence' in order to ground the precise concepts of formal arithmetic. Admittedly, the existence of non-standard models illustrates the difficulty of finding a suitable model-theoretic semantics in this area. (Though, the issues are complex. In the case of real numbers, the existence of non-standard models seems liberating, providing us with a deeper understanding of possible structures for analysis, and even new ways of doing mathematics in this area. Perhaps this is at least a cautionary example for the distinction between single-structure and 'algebraic' studies.) However, it might be argued, the last seventy years or so of proof-theoretic work has given us a rich body of technical knowledge of what follows from what in number theory.

The last point is surely true, but the significance of these studies needs further thought. There are two extreme positions that should probably be rejected. The first is that the 'metamathematics of arithmetic' is an arcane subject, sophisticated and technically accomplished, but in the end of little interest to the working mathematician, who use (however precisely in an informal sense) the informal notions. The second is a view of the kind gestured at in the previous paragraph: the informal concepts are 'folk' concepts, horse-and-buggy concepts to be replaced by the precise concepts developed in proof theory. The latter view disregards the fact that the axioms and rules needed to be intuitively sound, acceptable to the intuitions and practices of working mathematicians. The former view disregards the insights of proof theory, and its power to influence the 'working mathematician', both relatively directly, and through the growing cultural influence of computer science on mathematics.

Perhaps a more reasonable view is reached by adding a spin to the familiar 'reflective equilibrium' approach. The latter would suggest, for mathematics, something like the following picture. Growing experience with some structure, and particular judgments about it, suggests to mathematicians 'principles', such as the method of infinite descent, and that of mathematical induction. These principles in turn shape mathematicians' conceptions of the structures they are studying, and may affect particular judgments they make about them. We may add that there is no reason why such relatively course-grained principles of informal mathematics might not be sharpened up by the finer-grained concepts of proof theory. Similarly, the development of quantifiers and the study of recursive functions has enabled a deeper understanding of informal arithmetical thinking, and the articulation of a number of alternative frameworks (for example, finitary, constructive, predicative or set-theoretical) in terms of which to develop and justify formal 'versions' of informal arithmetic. Much of this work, for example on the power of PRA (primitive recursive arithmetic), is impressive and interesting in its own right, and might reasonably be expected to affect the way we think in, as well as about, informal arithmetic.(6)

What we seem to have, then, is a combination of, and interplay between the intuitive notion and various intuitively sound proof systems. We have been stressing the existence of a rich body of theory concerning the formal systems. It seems worth studying the intuitive notion further, to explore its contours, its strengths and weaknesses, and perhaps to look for other ways of explicating aspects of it that have not been captured so far.

5. *Exploring the intuitive notion.*

However, this is easier said than done. While it was assumed that the model-theoretic account provides a successful analysis or explication of the intuitive notion, there seemed little point in worrying too much about the latter. Even those who were not entirely happy with the standard story were inclined to work on modifying the technical concept rather than on investigating the intuitive one. Perhaps the latter could safely be consigned to the 'pre-his-

(6) Much of this material is reviewed in [Hajek and Pudlak 1993] as well as in [Sieg 1985], [Sieg 1988], [Simpson 1988] and in a series of very interesting papers by Solomon Feferman, e.g., [Feferman 1988].

tory' of the subject. As a result, perhaps, there seems to have been little serious investigation of the intuitive notion over the last fifty or more years. It is beyond the scope of this paper to do more than sketch a few of the issues that would need to be addressed in such a study.

5.1 *Culture and Cognition.*

Asssuming, then, that there is an intuitive notion of logical consequence, with the kind of significance that Etchemendy suggests, what sort of thing is it, and where does it come from?

Is it grounded in the social/cultural practices of argumentation in various forms? (And was Aristotle codifying, as it were, the structure of such practices?) If so, something like the picture suggested in the previous section would be quite plausible, with an ongoing interplay between more intuitive and more scientific conceptions. (So, perhaps would Wittgenstein's metaphor of town and suburbs.)

Or is it, somehow grounded in our cognitive system, perhaps in the architecture of a 'reasoning system', or in some other way? The development of cognitive science promises the possibility of understanding better the psychological basis of our reasoning skills without lapsing into the kind of 'psychologism' attacked by Frege and Husserl a century ago.

5.2 *Topic-neutral or Content-specific?*

Cutting across the first issue is one between logic as something that abstracts away completely from content, or as something that may involve, and even vary with, content. The Tarskian conception seems committed to the former view, despite Tarski's reservations about the possibility of characterizing the logical constants.

Etchemendy's account, on the other hand, suggests the idea that the logical consequence relation for a given language may reflect the meanings of expressions other than the familiar 'logical constants'. Recent work of Gardenfors and others on 'conceptual semantics' suggests one way in which a view on the previous issue could be combined with this perspective to develop, e.g., an account of the 'logic' of colour concepts. Harking back to the earlier discussion of arithmetic, Etchemendy's account suggests that what follows from what in the interpreted language of arithmetic is tied up in part with our number concepts. Understanding these, in turn, may involve reflecting on practices (e.g., counting) and cognitive structures.

5.3 *Naturalism and Modality.*

Perhaps the most perplexing issue, at least for a naturalist, concerns the link with modality. One of our strongest intuitions about logical consequence is: if a conclusion follows logically from some premises, then, if the premises are true, the conclusion must be true also. Etchemendy has an epistemic argument for this being required of any acceptable account of logical consequence. If he is right, we have another challenge for naturalism: how to give an account of logical truth and consequence that meets the modal requirement. Where does the intuition of necessity come from? And what in the world —or the mind— grounds it? We need to return to some of the older issues about logic, for despite the technical sophistication of model theory, and the undeniably valuable tools it has brought us, at a more fundamental level we may be as far from the 'goal of a scientific semantics' as when Tarski was writing in the 1930s. Or perhaps that goal was always an illusion?

Flinders University

References

Barwise, Jon, 1990, 'Consistency and Logical Consequence', in *Truth or Consequences: essays in honor of Nuel Belnap*, edited by J. Michael Dunn and Anil Gupta, Kluwer.

Barwise, Jon and John Etchemendy, 1989, 'Model-Theoretic Semantics', in *Foundations of Cognitive Science*, edited by Michael Posner, M.I.T.

Etchemendy, John, 1990, *The Concept of Logical Consequence*, Harvard.

Feferman, Solomon, 1988, 'Hilbert's program relativized: Proof-theoretical and foundational reductions', *Journal of Symbolic Logic* 53, 364-384.

Hajek, Petr and Pavel Pudlak, 1993, *The Metamathematics of First-order Arithmetic*, Springer.

Sieg, W., 1985, 'Fragments of Arithmetic', *Annals of Pure and Applied Arithmetic*, 28, 33-72.

Sieg, W., 1988, 'Hilbert's program sixty years later', *Journal of Symbolic Logic*, 53, 338-348.

Simpson, S., 1988, 'Partial realizations of Hilbert's program', *Journal of Symbolic Logic*, 53, 349-363.