What is Modeled by Truth in All Models?*

John MacFarlane†

March 23, 2000

Abstract

John Etchemendy has argued that the model-theoretic definition of logical truth fails as a conceptual analysis of that notion. I will argue that Etchemendy's criticism cuts deeper than recent critics have conceded. Properly understood, it is directed against the underlying analysis of logical truth as truth on all possible semantic interpretations of a language's nonlogical vocabulary, not against any particular mathematical realization of that analysis. The only way to block Etchemendy's argument is to reject his assumption that the possible semantic values for singular terms in an extensional language are the actually existing objects. In fact, a version of his argument goes through even if we retreat to the weaker assumption that the possible semantic values for singular terms are the possibly existing objects. I defend the model-theoretic analysis by arguing that there is a sense of "possible semantic value" for which both these assumptions are false.

1 Logical truth and truth in all models

It is commonly thought that the best precise account of logical truth we have is the one given in model theory: a sentence S is logically true just in case it is true in every model.[1] A *model* (for a first-order extensional language) consists in a nonempty set of objects—the *domain*—and an *interpretation function* that assigns

*Talk delivered at the Pacific APA, Albuquerque, NM, April 8, 2000.

†Department of Philosophy, University of California at Berkeley. E-mail: jgm@uclink.berkeley.edu.

[1]All of the points we will make in this paper about logical truth apply equally to logical consequence: we focus on logical truth solely because its model-theoretic definition is simpler.

objects in, functions on, and relations over this set to the non-logical singular terms, function terms, and relation terms of the language, respectively. Truth in a model (relative to an assignment of values to free variables) is defined recursively, in the manner of Tarski.

The model-theoretic account can be motivated as a strengthening of the traditional requirement that logical truths have no formal counterexamples.[2] A true sentence may fail to have counterexamples simply because the language happens not to contain non-logical terms with the requisite semantic values.[3] But whether a sentence is a logical truth should not depend on the poverty of its stock of non-logical terms: it should not be possible to turn a logical truth into a logically contingent sentence simply by adding new non-logical terms to the language. These considerations motivate a stronger requirement: logical truths must be true on every possible semantic reinterpretation of their non-logical terms (respecting semantic categories). The model-theoretic definition takes this *necessary* condition for logical truth to be a *sufficient* condition as well.

John Etchemendy [2] has argued at length that this necessary condition cannot be sufficient for logical truth. At the center of his case is an argument he dramatizes by considering a finitist metalogician. We will argue that although recent critics have been right to dismiss the "finitist" argument, they have dismissed it for the wrong reasons. The model-theoretic definition of logical truth *can* be vindicated, but not without addressing the question Etchemendy has forced us to ask: what is modeled by truth in all models?

2 Etchemendy's "finitist" argument

Let T be a sentence that is true in every *finite* model (i.e., every model with a finite domain) but false in some *infinite* model, for example:

$$\text{(T) } [\forall x\forall y\forall z(Rxy \,\&\, Ryz \supset Rxz) \,\&\, \forall x{\sim}Rxx] \supset \quad \exists x\forall y{\sim}Ryx$$

[2] That is, logical truths cannot be turned into falsehoods by uniformly substituting non-logical terms for non-logical terms of the same semantic category.

[3] For example, the sentence "some mammals are dogs" would lack counterexamples in a language whose only predicates were "mammal," "dog," "Dachsund," and "animal." (We must suppose in addition that the language does not contain a mechanism for the formation of logically complex predicates, like "non-dog.") Note that in an infinite domain there will be uncountably many possible predicate extensions, so (given the usual requirement that languages be effectively specifiable) there cannot be a predicate corresponding to each of them.

> (If R is a transitive, irreflexive relation, then R has a minimal element.)

A finitist, Etchemendy claims, can *consistently* assert both:

> (A1) There are only a finite number of objects, and
>
> (A2) T is not logically true.

However, (A1) is *inconsistent* with

> (A3) T is not true in all models.

For if there are only finitely many objects, then there are no infinite sets of objects, and *a fortiori* no infinite models (Etchemendy [2], p. 119). It follows that T is true in all models, since T is true in all finite models. Hence (A1) implies the negation of (A3). Since (A1) and (A2) are consistent, but (A1) and (A3) are inconsistent, (A2) and (A3) cannot be conceptually equivalent. Etchemendy concludes that there is more to being logically true than merely being true in all models.

It is important to notice that the point of this argument is *not* to cast doubt on the extensional adequacy, or even the *intensional* adequacy, of the model-theoretic account. Etchemendy assumes only that the finitist's claim (A1) is intelligible,[4] not that it is (or even *could be*) true: the argument goes through even if there are *necessarily* infinitely many objects ([2], p. 116). What the argument purports to show is that truth in all models is not a good *conceptual analysis* of logical truth.[5] Its target is

> . . . our assumption that the model-theoretic definition of consequence, unlike syntactic definitions, involves a more or less direct analysis of the consequence relation, and so its extensional adequacy, its 'completeness' and 'soundness,' is guaranteed on an intuitive or conceptual level, not by means of additional theorems. ([2], p. 4)[6]

[4]Some philosophers may question the very intelligibility of (A1), on the grounds that it does not make sense to count "objects" without specifying a sortal. But (A1) can be understood as: "there are only a finite number of objects of all sorts." Such a claim would be true if there were finitely many *basic* sortal concepts ($S_1 \ldots S_n$) such that (a) for each $1 \leq i \leq n$, the number of S_i's is finite, and (b) for every sortal concept S, every S is identical with an S_i for some $1 \leq i \leq n$.

[5]Etchemendy does not say what he *means* by "conceptual analysis," and in his argument he simply assumes that we have tolerably clear preformal notions of (conceptual) consistency and inconsistency. For the purposes of this paper, we will grant that these notions can be given a coherent sense. Note that in order to avoid the paradox of analysis, we must take these notions to concern the *norms* governing the use of concepts, not what is actually *thought* by someone who grasps them.

[6]In this paper we will be concerned only with Etchemendy's criticism of this view, not his

3 The triviality objection

Vann McGee agrees with Etchemendy that truth in all models is not a good conceptual analysis of logical truth ([7], p. 275). After all,

(B1) There are no sets

is consistent with

(B2) Not every sentence is logically true,

but inconsistent with

(B3) Not every sentence is true in all models.

For models are sets. If there are no sets, there are no models. And if there are no models, then it is *trivially* true that every sentence is true in all models. So (B1) implies the negation of (B3). Since (B2) and (B3) stand in different logical relations to (B1), they must not be conceptually equivalent. QED![7]

McGee's agreement is actually playful criticism. For if McGee's argument suffices to establish Etchemendy's conclusion, then Etchemendy is guilty of using a steamroller to squash a bug. Why worry about T, domain size, and the details of the model-theoretic definition of logical truth, if it is enough to point out that the definition presupposes that sets exist?

In effect, McGee's argument *trivializes* Etchemendy's claim that truth in all models is not a good conceptual analysis of logical truth. For if McGee's argument is cogent, then hardly *any* mathematically precise definition of an intuitive notion will count as a good conceptual analysis. Consider the inductive definition of the natural numbers as the smallest set that contains 0 and is closed under successor. Since it is consistent to hold that the natural numbers exist but sets do not, McGee's argument would rule out the definition as a conceptual analysis. But in fact, this definition is Etchemendy's *paradigm case* of a good conceptual analysis. Although Etchemendy concedes that "…it employs a variety of set-theoretical

attribution of the view to Tarski. For arguments that Tarski did not intend his definition as a conceptual analysis, see [3] and [8].

[7]We have taken the liberty of changing McGee's argument a little, in order to make it parallel to Etchemendy's own. McGee puts the argument as follows: "A model is a kind of set, and pure logic doesn't require that there exist any sets at all, so that it is logically possible that every sentence should be true in every model. But it is not logically possible for every sentence to be true, and, *a fortiori*, it is not logically possible for every sentence to be valid" (p. 275).

concepts that are not, by any stretch of the imagination, part of our ordinary understanding of the natural numbers," he maintains that it "...captures the essential feature of the intuitive notion..." ([2], p. 9), in a way that the model-theoretic definition of logical truth does *not*.

These considerations suggest that Etchemendy would reject McGee's simple argument against the model-theoretic definition of logical truth. However, McGee's argument is formally just the same as the version of Etchemendy's argument presented in section 2 above. They stand or fall together. Charity demands, then, that we find another formulation of Etchemendy's argument.

4 Models and analyses

So far we have been concerned with the relationship between logical truth and the mathematically precise notion of truth in all models. The key to a more charitable reading of Etchemendy's criticism is to see that there are really *three* notions in play here: logical truth, truth in all (set-theoretic) models, and truth on all possible semantic interpretations of the language's non-logical terms (henceforth, "PSIs")—that is, strong counterexample immunity. According to the view Etchemendy is opposing, the mathematically precise notion of truth in all models *models* the notion of truth in all PSIs,[8] which in turn is a *conceptual analysis* of the notion of logical truth:

$$\begin{gathered}\textbf{truth in all models}\\ \| \\ \text{(models)} \\ \Downarrow \\ \textbf{truth on all PSIs} \\ \| \\ \text{(is conceptually equivalent to)} \\ \Downarrow \\ \textbf{logical truth}\end{gathered}$$

What McGee's argument shows is that truth in all set-theoretic models is not conceptually equivalent to logical truth. This result—like the claim that model airplanes are not real airplanes—is indeed trivial, but Etchemendy is after bigger

[8] As Stewart Shapiro observes, the ambiguity of "model" makes this awkward to say ([10], pp. 138-9). See also Sánchez-Miguel in [9], p. 123.

game. He aims to show that the notion of *truth on all PSIs* is not conceptually equivalent to logical truth. Thus, he is interested in facts about set-theoretic models only to the extent that they represent features of PSIs.[9] Unfortunately, Etchemendy is not as explicit on this point as he might have been, and even those critics who do distinguish between truth in all set-theoretic models and truth in all PSIs have taken him to be concerned with the former.[10] Let us be more charitable.

5 Etchemendy's argument revised

Construed as an argument against the analysis of logical truth as truth on all PSIs, Etchemendy's argument runs as follows. Take T as before. A finitist can *consistently* assert both:

> (C1) There are only a finite number of objects, and
>
> (C2) T is not logically true.

However, Etchemendy claims, (C1) is *inconsistent* with

> (C3) T is not true on all PSIs.

Therefore, (C3) does not capture the essential conceptual content of (C2), and so "true on all PSIs" is not a good conceptual analysis of "logically true."

The crux of the argument is the claim that (C1) is inconsistent with (C3). In order to justify this claim, Etchemendy needs to say more about what is meant by "possible semantic interpretation" in (C3). Etchemendy considers two options:

> (R) *Representational semantics:* An R-PSI assigns to each non-logical term X the semantic value X would have had (given its actual meaning) in some possible world.
>
> (I) *Interpretational semantics:* An I-PSI assigns to each non-logical term X the semantic value X would have had in the actual world, given some possible meaning.

[9] He acknowledges that the underlying notion of truth in all PSIs can be modeled in a number of different mathematical frameworks, and he makes it clear that he is not concerned with any of these optional realizations in its own right. "Generally, the specific framework we presuppose is that of Zermelo-Fraenkel set theory, but of course nothing about the analysis dictates this particular choice, or even that our background theory should be a *set* theory rather than a *class* theory or *category* theory or *property* theory" ([2], p. 114); "…nothing about Tarski's analysis leads us to construe the variable *P* as ranging over sets rather than properties…" (p. 122).

[10] For example, Chihara ([1], p. 167) and Sánchez-Miguel ([9], pp. 120-123).

Representational semantics is clearly inappropriate for an analysis of *logical* truth: truth in all R-PSIs is just truth in all possible worlds, or *necessary* truth. Many sentences with formal counterexamples—for instance, "if Fido is a dog, then he is not a horse"—are true in all R-PSIs.[11] Etchemendy concludes that (C3) must be read as

(C3-I) T is not true on all I-PSIs.

But (C3-I) is demonstrably inconsistent with (C1). In order for (C3-I) to be true, T must be false on some I-PSI. T can be false on an I-PSI only if its quantifiers range over an infinite domain. But the elements over which the quantifiers range must be possible semantic values for singular terms.[12] So T can be false on an I-PSI only if there are infinitely many possible semantic values for singular terms. But in an extensional language, the semantic value of a singular term—that is, the way in which it contributes to the truth or falsity of sentences containing it—is simply the object it denotes. Since interpretational semantics considers only semantic values terms could have in the *actual* world, the possible semantic values for singular terms—no matter how their meanings are varied—are just the objects that exist in the actual world.[13] Hence, T can be false on an I-PSI only if there are actually infinitely many objects—which is precisely what (C1) denies. Thus (C1) and (C3-I) cannot both be true.

[11] It seems that in rejecting the model-theoretic definition, Etchemendy also rejects the narrow conception of logical truth that would exclude such sentences: see [2], p. 158, [1], p. 163, [8], p. 675, n. 63. However, Etchemendy does not regard truth in all R-PSIs as an *analysis* of logical truth ([2], p. 25).

[12] This assumption appears in the model-theoretic definition of logical truth as the stipulation that the interpretation function of a model can be *any* function that assigns elements of the model's domain to singular terms, sets of elements of the domain to predicates, etc.— so that if there is a model whose domain includes an object *m*, then there is a model that assigns *m* as the semantic value of some singular term. This amounts to the assumption that no item in the domain is *in principle* unnamable: it does not rule out the possibility that some objects may be unnamable because of our epistemic or pragmatic limitations. (Note that this is the *only* assumption about quantifier domains Etchemendy's "finitist" argument requires. The argument is insensitive to whether or not quantifier domains are relativized to interpretations, and to how the relativized domains are interpreted.)

[13] See [2], pp. 33-41. "Exist" here should be understood tenselessly: there is no implication that a singular term cannot denote, say, Shakespeare.

6 The combination approach

Given this construal of "PSI" in (C3), Etchemendy's argument looks unassailable. But must we accept this construal? The semantic value of an expression depends *both* on its meaning and on the state of the world. (The truth value of "there are winged dogs" can be changed, for instance, either by changing the meaning of "black" to "winged" or by genetically engineering winged dogs.) Each of Etchemendy's two readings of "PSI" considers possible variations in *one* of these factors, leaving the other fixed. But why not vary both, *combining* representational and interpretational semantics?

> (RI) *Representational-interpretational semantics:* An RI-PSI assigns to each non-logical term X the semantic value X would have had (given some possible meaning) in some possible world.

The combination approach—advocated by Stewart Shapiro and William Hanson[14] —is not subject to the objections we raised against representational semantics. It amounts to a further strengthening of the counterexample immunity demanded of logical truths: we now demand that logical truths remain free of counterexamples not just under expansions of the language, but also in counterfactual situations. Moreover, if we read (C3) as

> (C3-RI) T is not true on all RI-PSIs,

then Etchemendy's argument that (C1) and (C3) are inconsistent no longer goes through: the falsity of T on some RI-PSI does not imply the *actual* existence of an infinite number of objects.

But this response does not go deep enough: the analysis of logical truth as truth in all RI-PSIs is vulnerable to a modified version of Etchemendy's argument. Since the falsity of T on some RI-PSI implies the *possible* existence of an infinite number of objects, (C3-RI) is inconsistent with

> (C1-M) There *could not* be an infinite number of objects.

But (C1-M) is consistent with (C2), so (C3-RI) and (C2) are not conceptually equivalent.

[14] Shapiro [10], p. 148, Hanson [4], p. 388. On Shapiro's view, "Φ is a logical consequence of Γ if Φ holds in all possibilities under every interpretation of the non-logical terminology in which Γ holds."

The argument might be resisted by taking the modality invoked in (C1-M) and (RI) to be *logical* possibility.[15] So understood, (C1-M) is inconsistent with (C2): if it is *logically* impossible for there to be infinitely many objects, then T must be logically true. This move blocks the argument against the equivalence of (C2) and (C3-RI), but only at the cost of presupposing a notion of specifically *logical* possibility in the putative analysis of logical truth. Not only is this explaining the obscure through the more obscure,[16] but it threatens to be circular, for it is doubtful that our grasp of the specifically *logical* modalities is independent of the model-theoretic account of logical truth and consistency.[17]

No matter how the modality in (RI) is understood, then, logical truth cannot be appropriately analyzed as truth in all RI-PSIs. If the modality is non-logical, then the variant on Etchemendy's argument goes through. If the modality is logical, then the analysis threatens to be circular, or at best unilluminating.

7 "Possible semantic interpretation": an alternative construal

Must we then accept Etchemendy's criticisms? No. Consider the reasoning that generates the set of alternative construals of PSI we have been considering (R, I, RI). The reasoning starts from the premise

> (MW) The semantic value of an expression depends only on its meaning and the state of the world.[18]

Given this assumption, it seems to follow that there are only three ways of construing the modality in "possible semantic interpretation:"

- consider various possible worlds, keeping meanings fixed (R);
- consider various possible meanings, keeping the world fixed (I);

[15] This seems to be Shapiro's approach ([10], pp. 147-8). The modality in Hanson's version is not specifically logical ([4], pp. 383, 388).

[16] This is essentially Etchemendy's criticism of representational semantics as an *analysis* of the notion of logical truth ([2], p. 25).

[17] For an attempt to characterize logical necessity independently of logical truth, see [6].

[18] This claim is obviously false for languages with indexicals and intensional operators, in which other parameters—like speaker, time of utterance, and so on—may be relevant for the determination of semantic value. But it seems innocent enough for the simple extensional languages with which Etchemendy is concerned.

- consider various possible meanings in various possible worlds (RI).

We will argue that this reasoning is fallacious, and that there is a *fourth* sense of "PSI" which is *presupposed* by the three senses considered above.

The problem is not that (MW) is false—it is innocuous, if construed in its natural sense—but that it does not support the conclusion. The point is easier to see in the case of an analogous principle:

> (LL) The latitude and longitude of a ship depend only on its location on the surface of the earth.

True enough. *Given* the coordinate system that defines latitude and longitude, the ship's location determine its latitude and longitude. But if we did not already know this coordinate system, we could not determine the "possible latitudes and longitudes" by considering the ship's possible locations on the surface of the earth. (LL) *presupposes* a range of possible latitudes and longitudes: the range defined by the coordinate system itself. Similarly with (MW): the semantic value of a term is jointly determined by its meaning and the state of the world, but only against the background of a specification of the term's semantic category—that is, of the possible ways in which terms of that category can contribute to the truth or falsity of whole sentences. But this amounts to a specification of the range of possible semantic values in that category.

By way of illustration, consider Łukasiewicz's three-valued logic in which sentences that express future contingents receive the semantic value *i* (for "indeterminate").[19] *Given* Łukasiewicz's specification of the semantic category of sentences—that is, of the three different ways in which sentences can affect the truth values of more complex sentences of which they are components—we can intelligibly ask whether sentences could ever have *i* as a semantic value. Suppose that the world is deterministic—indeed, *necessarily* deterministic—so that there can be no future contingents. Then *no* combination of possible meaning and possible world will determine *i* as the semantic value of a sentence. Yet even under this supposition, there is a perfectly good sense in which *i* is a possible semantic value for a sentence: the sense determined by the semantic category itself. Indeed, unless one understands *this* sense of "possible semantic value," one cannot even *ask* which semantic values can result from various combinations of meaning and world. (In general, the question "which Fs are G" presupposes that we know what how to identify Fs and distinguish them from each other.) We suggest that it is

[19]See [5]. It is important here that *i* has a real interpretation; it is not merely an element in a matrix.

this sense of "possible semantic value"—and correlatively of "possible semantic interpretation"—that is invoked by the analysis of logical truth as truth in all PSIs. If we understand PSIs in this way, then the legitimacy of using *i* in PSIs for the Łukasiewicz logic does not depend on the possibility of a contingent future.

The same point holds for interpretations of first-order quantifier logic. In asking about the range of semantic values generated by variations in possible meanings and possible worlds, we *presuppose* a range of possible semantic values that depends on neither factor. That is, we presuppose a well-defined semantic category. It is not enough to say "objects" (just as, in the Łukasiewicz logic, it would not be enough to say "truth values"): we need to say what counts as an object and how objects are individuated. These questions are settled by our sortal concepts. Provided that our sortal concepts themselves do not *rule out* an infinite number of instances, there is a sense in which there can be an infinite number of possible semantic values for singular terms, even if there are *necessarily* a finite number of objects. If PSIs are assignments of possible semantic values in this sense, then Etchemendy's argument fails.[20]

It might be objected that these cases are not really analogous: whereas we can *model* the semantic values in Łukasiewicz's logic even if the universe is deterministic (using, say, the letters "*t*", "*f*", and "*i*"), we could not model an infinite domain if the universe were finite. But even if this is true,[21] it is irrelevant, since Etchemendy's argument concerns the underlying conceptual analysis, not the mathematical models. The fact that a finitist would have no good mathematical *model* of truth in all PSIs does not in any way impugn the analysis of logical truth as truth in all PSIs.

8 The semantic finitist

We can of course turn the crank once more and consider a *semantic* finitist, who rejects the idea that there are an infinite number of distinct possible ways in which a singular term might contribute to the truth value of a sentence. The semantic finitist argues directly (on semantic grounds) for

> (C1S) There is not an infinite number of possible semantic values for singular terms.

[20] Sánchez-Miguel makes a similar point in [9], pp. 119-124.

[21] In [1], Chihara argues that his "Constructibility Theory" can be used to model infinite domains without assuming that there are infinitely many objects (pp. 167-8).

But Etchemendy's argument does not go through unless (C1S) is consistent with

(C2) T is not logically true,

and we ought to deny this. The semantic finitist *ought to* accept T as a logical truth.[22] This result reveals a virtue, not a flaw, in the model-theoretic analysis: it is precisely because questions of logic and questions of semantics are linked in this way that logical truth is appropriately analyzed as truth in all possible semantic interpretations.

[22]Cf. [9], p. 121.

References

[1] Chihara, C., "Tarski's Thesis and the Ontology of Mathematics," pp. 157-172 in *The Philosophy of Mathematics Today*, edited by Matthias Schirn, Oxford University Press, Oxford, 1998.

[2] Etchemendy, J., *The Concept of Logical Consequence*, Harvard University Press, Cambridge, Mass., 1990.

[3] Gómez-Torrente, M., "Tarski on Logical Consequence," *Notre Dame Journal of Formal Logic*, vol. 37 (1996), pp. 125-151.

[4] Hanson, W. H., "The Concept of Logical Consequence," *Philosophical Review*, vol. 106 (1997), pp. 365-409.

[5] Łukasiewicz, J., "Philosophical Remarks on Many-Valued Systems of Propositional Logic," pp. 40-65 in *Polish Logic*, edited by S. McCall, Oxford University Press, Oxford, 1967.

[6] McFetridge, I. "Logical Necessity: Some Issues," pp. 135-154 in *Logical Necessity and other essays*, ed. J. Haldane and R. Scruton, Aristotelian Society Series, vol. 11, London, 1990.

[7] McGee, V., "Two Problems with Tarski's Theory of Consequence," *Proceedings of the Aristotelian Society*, n.s. vol. 92 (1992), pp. 273-292.

[8] Ray, G., "Logical Consequence: A Defense of Tarski," *Journal of Philosophical Logic*, vol. 25 (1996), 617-677.

[9] Sánchez-Miguel, M., "The Grounds for the Model-theoretic Account of the Logical Properties," *Notre Dame Journal of Formal Logic*, vol. 34 (1993), pp. 107-131.

[10] Shapiro, S., "Logical Consequence: Models and Modality," pp. 131-156 in *The Philosophy of Mathematics Today*, edited by Matthias Schirn, Oxford University Press, Oxford, 1998.