Notre Dame Journal of Formal Logic
Volume 34, Number 1, Winter 1993

The Grounds for the Model-theoretic Account of the Logical Properties

MANUEL GARCÍA-CARPINTERO SÁNCHEZ-MIGUEL

Abstract Quantificational accounts of logical truth and logical consequence aim to reduce these modal concepts to the nonmodal one of generality. A logical truth, for example, is said to be an instance of a "maximally general" statement, a statement whose terms other than variables are "logical constants." These accounts used to be the objects of severe criticism by philosophers like Ramsey and Wittgenstein. In recent work, Etchemendy has claimed that the currently standard model-theoretic account of the logical properties is a quantificational account and that it fails for reasons similar to the ones provided by Ramsey and Wittgenstein. He claims that it would fail even if it were propped up by a sensible account of what makes a term a logical constant. In this paper I examine to what extent the model-theoretic account is a quantificational one, and I defend it against Etchemendy's criticisms.

1 In earlier days of analytic philosophy, Frege and Russell defended an account of the logical properties (*logical truth* and *logical consequence*) that was soon severely criticized (and discredited, I would dare to say) by Wittgenstein and Ramsey.

Frege and Russell focused mainly on logical truth; they would have analyzed logical consequence in terms of logical truth. A logical truth, according to them, is either a plain truth characterized by being maximally general, or an instance thereof. This maximal generality lies in the fact that the only terms other than variables in a maximally general sentence (or *thought*, or *proposition*) are "logical constants": *and*, *or*, *not*, *for all*, etc.

This explanation has some intuitive plausibility to it. It seems that the truth of 'I went to the movies yesterday or I did not' is somehow general, independent of the facts about my visits to movie theaters. And it has the virtues of conceptual *reduction*: we eliminate a difficult modality (logical truth) in favor of something we understand much better: generality. But as Wittgenstein and Ramsey pointed out, it does not work. Wittgenstein put it in a characteristically elegant manner: the generality of the logically true sentences, he said, is not *accidental* generality.

Received September 12, 1991; revised October 6, 1992

Wittgenstein's and Ramsey's criticism is often presented as if it lay merely in putting forward some counterexamples. 'There are more than two things' is a typical one: assuming the first-order translation of it, and assuming too that the identity sign is a logical constant, it is a maximally general true sentence; but it is not an intuitive logical truth. But presenting it in this way, we miss the point of the criticism. And there is an obvious reply to the alleged counterexample from the Frege-Russell perspective. When we quantify in natural language, they could argue, we restrict our quantifiers to an implicitly specified domain. Therefore, the defender of the Frege-Russell view could argue, in giving the logical form of the sentence we should make that implicit assumption explicit. We would capture this logical form, thus, with something like:

$$(\exists x \in \mathbf{D})(\exists y \in \mathbf{D})(\exists z \in \mathbf{D})(x \neq y \wedge x \neq z \wedge y \neq z),$$

where '**D**' represents the intended domain of quantification. Now, '**D**' not being a logical constant, the sentence is not maximally general thereby. And the maximally general sentence of which it is an instance, namely,

$$\forall X((\exists x \in X)(\exists y \in X)(\exists z \in X)(x \neq y \wedge x \neq z \wedge y \neq z)),$$

is simply not true.

I believe that this kind of answer is always available to someone defending the Frege-Russell account. The reason will be clear, I hope, at the end of the paper. We could say that the account is extensionally correct, if it were not for the vagueness implicit in the concept of *logical constant*, which makes it difficult to speak about extensions. The problem with the account, and the point of Wittgenstein's and Ramsey's criticism, is a different one. The problem is that it is no account at all. In spite of its intuitive plausibility, there is no explanatory connection between the ordinary concept of generality and the concept of logical truth. *Logical truth* is a species of *necessary truth*; but a general truth, in the ordinary sense of 'general', need not be a necessary truth. And if it is alleged that the generality in question is one of a special kind, an account should be given of its specificity.

This difficulty is obvious when we pass from *logical truth* to *logical consequence*, the concept we most want to understand—and the one the Frege-Russell account places in the background. The account has it that a valid argument, like 'every Greek is mortal, Socrates is Greek, therefore Socrates is mortal', is valid because it exemplifies a valid scheme of inference. There are infinitely many valid schemes of inference, but we can manage to reduce all of them to the repeated application of a few of them, perhaps just one. But we still need an explanation of why the primitive schemes of inference, in their turn, are valid. This is what the account cannot provide. Assuming that the account explains the logical validity of maximally general truths (which it does not), it is immediate how it accounts for the logical validity of their instances. But this immediacy disappears when it comes to explain the logical validity of a schema of inference in terms of the truth, or even the logical truth, of any maximally general sentence. We can almost hear Carroll's tortoise asking the embarrassing question: "Granted, *for every concept X and Y and for every object x, if every X is Y and x is X, then x is Y* is true; I even grant that, being maximally general, it is a logical truth. But what has that to do with the validity of *inferring* a proposition with the form *x is Y* from two propositions with the forms *every X is Y* and *x is X*?"[1]

This becomes even more embarrassing when we realize that the reason why the schemes are valid must be the same as the reason why their instances are. (Assume that the only primitive rule of inference is a general scheme of instantiation: what was operating in the example above was just a particular case of it.) The scheme allows us to "explain" why their instances are logically valid arguments. But we lack an explanation of the validity of the scheme itself. And that shows that the explanation was faulty, because intuitively the same explanation must be given for the validity of the scheme as for that of their instances. In this respect, the Frege-Russell account is not better off than the several "syntactic" accounts, popular in the early thirties, according to which a logical truth is a sentence whose "logical form" is an axiom or a theorem of some formal calculus.

Be that as it may, the bleak prospects for the Frege-Russell account do not upset us very much because we think we have an alternative account that we trust, an account that strikes us as obviously correct: the model-theoretic account that has its roots in the work of Alfred Tarski.[2] And this near unanimity is rare in the philosophical domain.

Alas Etchemendy [6] aims to establish "that the standard, semantic account of logical consequence is mistaken" (p. 8). Etchemendy claims "that Tarski's analysis is wrong, that his account of logical truth and logical consequence does not capture, or even come close to capturing, any pretheoretic conception of the logical properties" (p. 6). And it is clear from the context that by 'Tarski's account' Etchemendy does not mean here "the account that, as a matter of historical fact, Tarski proposed," but simply the model-theoretic account many of us trust, according to which "a sentence is logically true if it is true in all models; an argument is logically valid, its conclusion a consequence of its premises, if the conclusion is true in every model in which all the premises are true." (p. 1).

The reasons Etchemendy gives in favor of his contention are essentially Wittgenstein's and Ramsey's reasons to reject the Russell-Frege view. And I think his criticism is well taken, because many people seem to understand the model-theoretic account as just a variant of that view. Tarski himself, in certain passages of "On the Concept of Logical Consequence" (the paper many consider the foundation of the model-theoretic approach) seems to understand it that way.

Nevertheless, I am going to contend in this article that Etchemendy's main point is mistaken. After presenting Etchemendy's views (Section 2), I will argue that there is something more in the standard, semantic account of logical consequence than the Frege-Russell idea (Section 3), and that the account is correct, at least, when restricted to languages or fragments of languages that are "first-order translatable" (Section 4). I will also argue that it is to some extent reasonable to suppose that a view like that could have been "what Tarski had in mind" in writing certain passages of "On the Concept of Logical Consequence" (Section 5), although I do not want to dwell on the historical point too emphatically.

2 Etchemendy bases his criticism on a distinction he finds natural to draw between two different ways of viewing the models of the model-theoretic account; he labels these two viewpoints *representational* and *interpretational* semantics. In representational semantics, models are viewed as representing possible ways the world could have been. In interpretational semantics, models represent possible "interpretations" or "meanings" the nonlogical signs of the language could have had. Etchemendy stresses that the difference is subtle but important. Etchemendy's argument develops from the (correct) assumption that the model-theoretic account of logical properties presupposes the interpretational view of models.

It is important to bear in mind why that assumption is correct, why what Etchemendy calls "representational semantics" has nothing to do with the project of explaining the logical properties. When we set out to find an explanation of the logical properties, what we are looking for is an understanding of the "modal knot" that is a constitutive part of these properties. Logical truths *must* be true; if the premises of a logically valid argument are true, the conclusion *must* be true. But in doing representational semantics, we are taking the logical space (the space of all possible worlds) as given: we rely on our unexplained intuitions about which worlds are possible to determine which truths are logically true. According to the representational view of models, the question "Is 'John is not father of himself' a logical truth?" reduces to the question "Is there a possible world in which John is his own father?" Obviously, this is not the kind of reduction from which we should expect a great amount of illumination.

Let me now summarize Etchemendy's presentation of the model-theoretic account, and his reasons against the account. For ease of exposition, let me consider from now on, except when I explicitly say the contrary, only the concept of logical truth—as Etchemendy himself does.

Let L be the language of whose logical properties we are interested in providing an account.[3] We start by selecting a subclass from the vocabulary of L, the class F of the *fixed terms* or *logical constants* of L. Then the model-theoretic account, viewing models interpretationally, consists of the following contentions: a sentence σ of L is *logically true* (with respect to our selection of F) if and only if σ is just *true* relative to any *semantically reasonable* interpretation of the nonlogical terms in σ. A sentence σ of L is a *logical consequence* of a set Γ of sentences of L if and only if σ is true relative to every *semantically reasonable* interpretation of the nonlogical terms in $\Gamma \cup \{\sigma\}$ such that relative to it every sentence in Γ is true.

An interpretation for the nonlogical terms of L is just a *model* for L. An interpretation is *semantically reasonable* if, for instance, the interpretations of predicates are subsets of the domain of quantification, which is determined by the interpretation of quantifiers,[4] and the interpretations of constants are members of the domain.

Etchemendy then claims that, on this *interpretational* view, the model-theoretic account is a *quantificational* account of logical truth (as the Frege-Russell account is). Let $\sigma(c_1, \ldots c_n)$ be a sentence of L, with $c_1, \ldots c_n$ its nonfixed terms. Let $\sigma(x_1, \ldots x_n)$ be the sentential function resulting from $\sigma(c_1, \ldots c_n)$ by substituting appropriate free variables (possibly predicate variables) for its nonfixed terms. Then the original sentence is logically true, according to the account, if and only if the formula

$$\forall x_1, \ldots x_n \sigma(x_1, \ldots x_n)^5$$

is just true.

In this way, the model-theoretic account—on the interpretational view of models – tries to reduce logical truth to plain truth. Etchemendy calls the principle embodied in the model-theoretic account the Reduction Principle: it states that a sentence is logically true if and only if it is an instance of a true universally quantified sentence whose constant terms belong to F (the set of logical constants of L).[6] Speaking carelessly, the model-theoretic account reduces the logical truth of a sentence σ to the plain truth of something like "for each possible meaning of the nonlogical signs in σ, σ is true."

I have said "speaking carelessly" because there are obvious difficulties in taking literally the word 'meaning' in the foregoing reduction. First of all, how can we take seriously the idea that a truth value could be the interpretation of a sentence, in the usual sense of the word 'interpretation'? And the idea that a set could be the interpretation of a predicate is not more plausible. Second, there is also the problem of accommodating the constraint that the interpretations of the nonlogical terms must agree with each other – a constraint hidden in my description beneath the idea that the interpretations should be "semantically reasonable." In Chapter 5 of [6] Etchemendy wonders whether this constraint on interpretations is coherent with what he takes to be the interpretational view of the model-theoretic account, and I think he is right in pointing out that here there is a problem for it.

To reach the conclusion that the model-theoretic account relies on the *Reduction Principle*, Etchemendy has followed, faithfully it seems, Tarski's argument in "On the Concept of Logical Consequence" (see [2]). Tarski starts by considering approvingly the *substitutional* account of the logical properties: a logical truth is a truth that remains always so, when the nonlogical terms in it are substituted for by others belonging to the appropriate category. He claims that the substitutional account has two important virtues: first, it captures both what he calls the "formality" and also the modality present in the intuitive concept of consequence; and second, it avoids a fatal shortcoming of the substitutional account, namely, the dependence of the output on the expressive power of the language.[7] The problem lies in that, in the pure substitutional account, if 'George Bush' were the only singular term of the language, the account would force the result that 'George Bush is President of the U.S.' is a logical truth. And then he proposes the model-theoretic account as a way of keeping the two virtues while avoiding the shortcoming.

Tarski seems to be claiming, then, that 'everything is a dog' is not a logical truth, *because* the relevant general sentence, 'for every possible meaning of 'thing', X [i.e., for every possible domain], and for every possible meaning of 'is a dog', Y, every X is Y', is not true; and 'George Bush is President of the U.S.' is not a logical truth because the relevant general sentence 'for every possible meaning of 'George Bush', x, and for every possible meaning of 'is President of the U.S.', Y, and for every possible meaning of (the implicit word) 'thing', X [i.e., once again, for every possible domain], x is a Y among the X' is not true. But then, it is difficult, as Etchemendy points out, to explain the rationale of imposing restrictions on what "meanings" we can consider for certain nonlogical expressions, given what other "meanings" we are currently considering for others. The main thrust behind the account, which seems to be the same as that motivating the Frege-Russell account, is apparently lost when we take into consideration restrictions on the "meanings" we can assign to a given word, depending on the "meanings" we had already given to some other words. (To give one example of an apparently unjustified restriction, the "meanings" we assign to the predicates must be subsets of the "meaning" we had provided for the expression indicating the domain.) The model-theoretic account, as derived from the substitutional one, was supposed to be that a sentence is logically true *because* it stays true whenever we vary the meanings of the nonlogical terms in it; and the only restrictions which seem justified here are those imposed by the semantic category of the expression.[8]

The main objection Etchemendy levels against the account focuses on the weaknesses of the *Reduction Principle* embodied in the model-theoretic account. There are, as a matter of fact, two variants of the objection, corresponding to two different explanatory goals (one more ambitious than the other) that may be attributed to the model-theoretic account; I will be concerned with only the modest version.

Taken with the more ambitious goals, the model-theoretic account attempts to offer a definition of the general concept of *analytic truth*—truth in virtue only of the meaning of certain expressions. In that case, the account is clearly wrong, because *being an instance of an universally quantified sentence whose terms other than the variables are all among those in* F is a necessary but not a sufficient condition of *being true in virtue only of the meanings of the terms in* F. For instance, the fact that, whatever the domain, and whatever the "meaning" of 'John' in it, 'John is not a father of himself' is true should not be sufficient to say that that sentence is true in virtue only of the meaning of 'is a father of'. In other words, the fact that 'John is not a father of himself' is an instance of 'for every domain X, and for every x in X, x is not a father of x', a true sentence whose only terms other than variables are 'is a father of', 'not' and 'for every', should not suffice for the sentence being true in virtue of the meaning of 'is a father of' (plus that of 'not' and 'for every'). Otherwise, many factual truths would be accounted as true in virtue of meaning. This itself could be an example; or we could find others, for those who think that 'John is not a father of himself' is in fact true in virtue of the meaning of 'is a father of'.

As I said, the second version attributes more modest goals to the model-theoretic account, and it is the one that concerns us here. According to it, the model-theoretic account does not pretend to explain the general concept of *truth in virtue of meaning*, but only the particular concept of *truth in virtue of the meaning of the "logical constants."* The account thus relies on a previous explanation of what makes an expression a *logical constant*. (Such an explanation is lacking in the Frege-Russell account and in what we have said so far about the model-theoretic account.)

Now, whatever the explanation of the nature of the logical constants, Etchemendy contends, the account would also be wrong if intended with these more modest goals. Here is his argument: there are sentences that would have had the distinctive mark of the logical truths (i.e., would have been instances of a generally quantified true sentence whose terms other than the variables are all logical constants), and so would have counted as logically true by the account, *against our intuitions*—if we had not made an extralogical assumption. This allegedly shows the intuitive inaccuracy of the account. It is an essential part of our intuitive understanding of the concept that logical truths do not depend on matters of contingent or otherwise substantive fact. But whether or not a sentence has the property of being true in every model depends on the correctness of substantive, extralogical assumptions about the existence of models.

Consider the example I discussed before, 'there are more than two things'. I said that the defender of the Frege-Russell account has an easy answer against the claim that this is a counterexample to his analysis. The answer was that the logical form of the sentence must include a nonlogical term referring to the intended domain of quantification. It must be something like

$$(\exists x \in \mathbf{D})(\exists y \in \mathbf{D})(\exists z \in \mathbf{D})(x \neq y \wedge x \neq z \wedge y \neq z),$$

and then the maximally general sentence of which this is an instance, namely,

$$\forall X((\exists x \in X)(\exists y \in X)(\exists z \in X)(x \neq y \wedge x \neq z \wedge y \neq z))$$

is simply not true.

Etchemendy's argument can be seen as another turn of the screw. Why is that sentence not true? Because there are domains with less than three objects. But this is an extralogical assumption. If, as a matter of fact, there were only two objects in the universe, the account would indeed *overgenerate*: it would count as logically true sentences which intuitively are not so (in this case, the negation of the sentence under consideration).

Etchemendy's examples are slightly different, although the point of them is the same. Suppose that there were only a finite number of objects, say n. Then each of the sentences that say that there are fewer than $n + m$ objects, for each m greater than zero, would be declared by the account logically true. If those sentences are not so declared, that has nothing to do with the account itself; it is only because in the background set theory we make an assumption to the effect that there are domains with an infinite number of objects. But this is a substantive, extralogical assumption. Etchemendy claims that it is in the second-order case that the account more obviously fails, because those set-theoretic background assumptions force the account to overgenerate alleged logical truths.[9]

The argument by Etchemendy summed up here is fallacious. I think that even a simple quantificational account, like the Frege-Russell account, has an acceptable answer to it. The reply in both cases is that what is in question is not what *actual* domains of quantification there are, but what *possible* domains of quantification there are, and that is not a substantive or otherwise nonlogical matter. In any event, I do not intend to defend the Frege-Russell account (which I believe misguided on independent grounds); and it would be difficult to do so, precisely because of those deficiencies that are to be held responsible for what I have described before as its lack of explanatory power. Regarding the model-theoretic account, however, that reply to Etchemendy's criticism is exactly right. To show that, I need to examine carefully that account and its theoretical underpinnings. I thus leave the discussion of what I reckon as Etchemendy's fallacy until I have introduced what I believe is lacking in his understanding of the model-theoretic account.

3 The *Reduction Principle* embodied in the model-theoretic account of *logical truth* is a consequence of a partial semantic theory for the language under study, and it is as a consequence of that fragmentary semantic theory that it is correct to say that a logical truth is a truth in all models, or "interpretations," if we prefer that confusing word. Let me explain.

It is no mystery why any assertion made by uttering the sentence 'if today is Monday, then tomorrow is Tuesday' *must* be true.[10] The conditions that any such assertion imposes on the world for it to be true are already fulfilled, given the meanings of 'today', 'tomorrow', 'if . . . then', 'Monday' and 'Tuesday'. A good semantic theory for English must have it as a consequence. And it must also have as consequence that to assert 'tomorrow is Tuesday' as an inference from an assertion of 'today is Monday' is to infer soundly.[11]

The idea behind the model-theoretic account, as many people have said before, is that *logical truth* and *logical consequence* are species of *analytic truth* and *analytic consequence*; that logically valid inferences are analytical inferences like the one shown before, except that they depend on the meaning of a very particular set of expressions: 'every', 'there is', 'and', 'if . . . then', etc. What makes these expressions specific is hard to say. Whatever it is, it should account for something that makes the former analogy weak in one certain respect: whereas 'today', 'tomorrow', 'Monday', 'Tuesday', are not present in every language (they are not present, for instance, in the fragment we use in doing mathematics), the paradigm cases of "logical constants" are much more widespread. Arguably, some of them must be part of anything we would like to call *language*. Any account of the specificity of the "logical constants" should then explain why logically valid arguments do not seem "conventional" in the way my former example does. In any case, the analogy is good in that it shows why, if *logical truth* is *truth in virtue of meaning*, we thereby understand where the "modal knot" in the logical properties come from.[12]

Anyway, none of this should concern us here; I do not need to explain what it is that makes the logical constants peculiar, that semantic feature which philosophers have tried to capture as their "contentless" or "formal" character. There are a variety of interesting ideas in the literature that could be explored, including Hacking's explanation [7] that logical constants are 'a mere "subproduct" of the system of representation' by characterizing them as being given by the introduction of elimination rules of sequent calculi meeting certain constraints, and their meanings obtained from those rules, and Sher's recent development [12] of Tarski's and Dummett's idea that logical constants are those *invariant under permutations of the universe*. I will take for granted that the ones corresponding to the "fixed terms" of first-order logic *are* logical constants.

Now, as Etchemendy says, if we consider the model-theoretic account as purporting a general characterization of the sentences true in virtue of the meaning of the expressions in class F, for arbitrary F, it is clearly mistaken. The model-theoretic account has, however, more modest goals; it attempts only to characterize *truth* (and *consequence*) *in virtue of the meanings of the logical constants*. When models enter the picture, their role in the account must be required by a previous hypothesis about the meanings of the logical terms. Thus, according to my interpretation, the model-theoretic account must (and does, in fact) involve a semantic theory for the logical particles. This is what was lacking in the pure quantificational accounts of Frege and Russell. But in fact, when we give the semantics for a first-order language, and we give the semantics of the logical or fixed expressions, we are making hypotheses about the meaning of these expressions in the fragments of language about whose logical properties we are theorizing. This partial semantic theory is as fundamental a part of the account as the well known definition in terms of models. More fundamental, indeed, because it is what gives sense to the definition. Let us see how.

That partial semantic theory tells us that, syntactically, the logical expressions contribute to the formation of complex sentences starting from the "atomic" ones (not really sentences, but "protosentences" or sentential functions in general); and, semantically, they determine the truth conditions of complex sentences given certain semantic properties of other less complex ones. And here is the point that interests us. For, although the complete truth conditions of, say, an atomic sentence's negation will depend on richer semantic properties of the sentence than its truth value (perhaps they involve a particular and a property), the contribution to those truth conditions made by the negation sign itself depends only on the sentence's truth value (assuming the theory captures correctly the semantics of negation). And although the entire truth conditions of ⌜for every β δ⌝ depend on richer semantic properties of protosentences β and δ (assume they are atomic) than the class of those entities in our intended domain which are β and the class of those which are δ, the contribution to them made by 'every' depends only on these semantic properties of β and δ. Whereas to give the meaning of the expressions occurring in atomic sentences you must, among other things, refer to objects-in-the-world and properties of them, to give the semantics of the logical expressions you need only mention certain abstract semantic properties of assertions and components thereof; you need only mention, for instance, the truth or falsehood of assertions, or the truth or falsehood of something in some previously specified domain of parts of assertions ('propositional functions').[13]

Now it is *relative to that partial semantic theory*, that to say that a sentence is a truth in all "models" or "interpretations," in a certain sense of those words that I will explain presently, is just to say that it is a logical truth (a truth in virtue of the meanings of the truth-functional connectives and the universal and existential quantifiers). Let me use the term 'logical value' to refer to the semantic properties of the nonlogical expressions to which alone the truth-conditional contribution of the logical expressions is sensitive, according to the meanings that our semantic theory attributes to them. The *logical value* of a sentence, for instance, (in the first-order case) is its truth value, and the *logical value* of a sentential function is the collection of individuals in some specified domain of which it is true. The intended domain of quantification is also a *logical value*. Although, strictly speaking, it is not the value of any nonlogical expression, but only implicitly taken for granted in the conversational context, following Etchemendy we could think of it as the logical value of the '-thing' part of the ordinary quantifiers. Last, let us define a *preformal model* as a *possible set of logical values such that expressions belonging to the same logical category as the nonlogical expressions in the sentence or argument could have those values.*

I would like to emphasize that preformal models, according to this understanding, are not "possible worlds," something like possible, maybe counterfactual, extensions that the sentential functions could have had in other possible circumstances, keeping the concepts they express fixed. Nor are models "possible meanings" that the nonlogical expressions could have had. In doing logic, we do not care about the meaning of the nonlogical expressions, only about their logical value; or, to put it accurately, only about the logical value that *expressions with that logical syntax* could have. What we call 'interpretations' of nonlogical terms when we do logic are not really interpretations. They are only either the actual semantic properties of these expressions to which, according to our semantics, the logical expressions under study are sensitive, or semantic properties with the same characteristics that the nonlogical expressions, given their logical category, could have. When we say that 'it is not the case that this patch is entirely red or this patch is entirely green' is not a logical truth (a truth in virtue of the meaning of 'it is not the case that' and 'or'), because there is an interpretation of 'this patch is entirely red' and 'this patch is entirely green' in which the sentence is false, we are not claiming either that those two same sentences are actually false or that there is a possible world in which those two sentences, keeping fixed their meanings, are false. We are claiming only that two different atomic sentences could both be false. To repeat: preformal models are possible logical values that expressions belonging to the same logical category as the nonlogical expressions in the sentence could have had. They represent possible combinations of those semantic values of the nonlogical expressions on which the specific contribution made by the logical expression to the truth conditions of the whole depends.

Now I claim that *being true in virtue of the meaning of the logical constants* and *being true in all preformal models* are one and the same property. To see this, let me discuss two examples before offering a more general justification. Consider first a simple case. We want to know whether 'this patch is red or this patch is not red' is a logical truth, and why. To answer this question, we have a semantic theory that tells us that the meanings of the disjunction sign and that of the negation sign (in English) are given by their truth tables. Our semantic theory says also that the logical structure of our sentence is: $(p \vee \neg p)$. Is this a truth only in virtue of the meanings of 'or' and 'not'? The contribution to the truth conditions of the sentence made by these two expressions, according to what our semantic theory says their meanings are, depends, in specific ways, on the truth value of 'this patch is red'—*and it does not depend on anything more.* Our semantic theory cannot give us the truth value of that sentence, because that depends on nonsemantic matters. But if, for every possible truth value that a sentence could have (that is, *in every possible preformal model*), our original sentence is going to be true (as it is in this case), we can be sure that it is true only in virtue of the meanings of the two logical expressions. And, on the other hand, if there is a preformal model in which the sentence is false, then knowing the meaning of the logical expressions in it alone does not determine whether it is true. More generally, any assertion constructed from atomic assertions only by means of disjunctions and negations that is true in every model (that is, in every assignment of truth values to the atomic assertions) must be true only in virtue of the meaning of 'or' and 'not'—if the assumed semantic analysis of 'or' and 'not' is correct, and if by 'model' we understand what I explained above.

Consider one more case. Is 'every father is parent of some child' a logical truth? Assume that we can correctly capture the "logical syntax" of this sentence by means of the first-order sentence '$\forall x(F(x) \rightarrow \exists y(C(y) \wedge P(x,y)))$'. The assertion, if our semantics for the logical expressions is correct, must be about some particular domain, and the sentential functions 'being a father', 'being a child' and 'being parent of' will have particular extensions in that domain. The contribution to the truth conditions of this sentence made by the logical expressions depends only upon them. The extension of a sentential function is a contingent semantic property of it; that is, in general, the semantic theory alone cannot give us the extension of a sentential function nor, as a result, the truth value of our sentence. But we are asking whether or not the assertion is true *only in virtue of the meaning of the logical expressions*. And the semantic theory gives us what we need to establish it: what is a possible domain of quantification, what is a possible logical value for nonlogical expressions of the same syntactic category as the nonlogical expressions in our sentence, and the meaning of the logical expressions in it. So, if the sentence is true in all models, in the previously defined sense of 'model', then we can be sure that it is true only in virtue of the meaning of the logical expressions. And if not, as in this case, it cannot be a truth only in virtue of the meanings of the logical expressions, because knowing their semantics alone does not establish the truth of the sentence.

The foregoing was by way of example. Let me now give a more general argument in favor of my claim, namely the following biconditional: a sentence is logically true, in the intuitive sense of the notion, if and only if it is true in all preformal models.

($\Rightarrow$) If the truth of a sentence follows from the semantics of the logical expressions in it, then the sentence is true in all preformal models. Let σ be a sentence that is not true in all preformal models. Now, one of the following two things could happen: (i) Given their (full-fledged) meanings, there is a possible world in which the nonlogical expressions in the sentence have the logical values that they have in one of those models which witness that it was not true in all models. But it is a contingent matter (that is, not a purely semantic matter) that the nonlogical expressions do not have, as a matter of fact, precisely those logical values in the actual world; so the truth value of the sentence could not have been decided, after all, only by means of semantic theory. Or (ii) there is no semantically possible world which corresponds to any of the models that showed that the sentence was not true in all models (that is what happens in the second of my previous two examples, 'every father is parent of some child', as the reader surely realized.) Then the sentence is indeed an analytic truth, but it is not true only in virtue of the meanings of the logical constants. It is true (in part, at least) in virtue of the meanings of some of the nonlogical expressions in it.

($\Leftarrow$) If a sentence is true in all models, we can establish that fact by appealing only to the part of our semantic theory that concerns the logical constants. Why? We can put it in this way. Suppose you know the meanings of '$\forall$' and '$\rightarrow$', these meanings being what the truth definition for a first-order language says they are. Consider a sentence with the logical structure represented by '$\forall x(P(x) \rightarrow P(x))$'. Having recourse only to your semantic knowledge, usually you cannot know either what the intended domain of quantification is or what the extension of '$P(x)$' in that domain is. But you know, having recourse only to your semantic knowledge, what a possible domain of quantification is, and also what a possible extension for a sentential function like '$P(x)$' is. With that knowledge, you can establish that the sentence is true, whatever the domain of quantification and the extension of the sentential function in it (i.e., it is true in all preformal models). On the other hand, it would not be true, if, say, 'every' meant what 'no' means. So it is a truth in virtue of the meanings of the logical particles in it, and in virtue of nothing but these meanings. Therefore it is a logical truth in the intuitive sense of the word. (I have appealed to the idea of semantic knowledge to make things more easily "surveyable," but I would rather put it in terms of what follows from an adequate semantic theory for the logical expressions.)

The reader may have sensed a weakness in this last part of the argument; he could have asked himself to what extent the notion of *preformal model* gives us anything close to what we expect of a concept of *logical consequence*. Of course, this notion cannot be pushed too far. But I want to stress here that I do not need to do so (as it will become clear in the next section, where I discuss Etchemendy's argument). To make my point I need only claim that the full-fledged models of the model-theoretic account, which are precisely characterized with the help of set theory, may be seen as developments of their preformal counterparts. The formal notion of model cannot conflict with the preformal one where our intuitions about logical truth and consequence are clear; but it can indeed go further, and it does. Going further than the intuitive one, the precise model-theoretic account, given for a formal language and employing a formal notion of model, not only *explains* the intuitive concept of consequence, but to a considerable measure *shapes* it. Nevertheless, this is not surprising. Exactly the same happens, for instance, with any good formal account of *grammaticality*.

To sum up. We have preformal concepts of the logical properties, applying to sentences and arguments in natural language. We build formal counterparts of fragments of natural languages (and, of course, formal languages without counterpart in natural languages too, but those do not concern us here), distinguish fixed terms from other, nonlogical terms whose interpretation varies from structure to structure, and define in the usual way the formal counterparts of the preformal concepts of the logical properties. To the extent that these formal structures are a good characterization of what I have labeled *preformal models*, and that the formal semantics for the logical constants corresponds reasonably well to the actual semantics of their counterparts in natural language, the model-theoretic account is an intuitively correct explanation of the preformal concepts of the logical properties, because it comes to say that a logical truth is a truth in virtue of the meaning of the logical constants (and explains why).[14]

The main difference between the model-theoretic account and any purely quantificational account, like the Frege-Russell account or the substitutional account, is that the former presupposes and includes an account of the semantics of the logical particles. On the basis of that account, and only on the basis of it, it is claimed that the truth of a general sentence ("σ is true in every model") is not only necessary (the left-to-right direction of the biconditional above, undisputed by Etchemendy), but also sufficient (the right-to-left direction that Etchemendy disputes) for the logical truth of another (σ). Notice that the cross-term restrictions on *interpretations* that Etchemendy questions are perfectly well justified on my understanding of the model-theoretic account. The same semantic theory that gives us the *possible logical values* we need to combine tells us what restrictions there are on the possible logical values for some expressions, even relative to the logical values of others.

As I said before, in the *Tractatus*, criticizing Russell's explanation of the concept of *logical truth*, Wittgenstein says that generality is not the mark of logical truths, that a universally quantified sentence could be true only accidentally, not for logical reasons. But he also adds that it could be said that logically true sentences are universally true "in an essential, as opposed to accidental, sense of 'universal'" (Wittgenstein [14], §§6.1231 and 6.1232). With the help of the foregoing discussion, we can make sense of the *essential generality* of logical truth hinted at here by Wittgenstein. When we explain why 'this patch is red or this patch is not red' is a logical truth, we have explained why every sentence with the same logical structure is a logical truth. Remember what models represent, according to my account: possible logical values that *expressions belonging to the same category as the nonlogical expressions in the sentence could have*. They do not represent the logical values of the nonlogical expressions in the sentence, nor possible logical values these expressions could have had in other possible circumstances, keeping their meanings fixed. Besides, we can make sense of the old idea that logical truths are truths "in virtue of the form alone," (an idea that Etchemendy criticized in [4]). This cannot mean that logical truths are true only in virtue of syntactic facts (even in a very abstract sense of 'syntax'). They are truths in virtue of semantic facts; the meaning of the logical constants are semantic properties, and of course the concepts of *truth* and *true of* are semantic concepts. But their truth depends only on the *logical syntax* of the nonlogical expressions, not on their *content*.

4 As Etchemendy stresses in criticizing the ambitious version of the model theoretic account, it would be a mistake to claim that, say, the sentence 'if Albert Gore is a U.S. senator, then Albert Gore has a sister' is true in virtue of the meaning of 'U.S. senator', 'having a sister' and 'if . . . then' if and only if the sentence 'for every x, if x is a U.S. senator then x has a sister' is true, intending that the right-hand side of the biconditional give an analysis of the left-hand side, even if the biconditional happens to be materially correct. A way to see this is to observe that whether or not the sentence 'if Albert Gore is a U.S. senator, then Albert Gore has a sister' is true in virtue of meaning is a conceptual fact, whereas whether or not the sentence 'for every x, if x is a U.S. senator then x has a sister' is true is a contingent matter. *Analysans* and *analysandum* have clearly got different properties.

Etchemendy claims that the same criticism applies to the model-theoretic account. Even when it is materially correct (as it arguably is in the first-order case) it is incorrect *as an account*, because, whereas the contention that a given argument is logically valid is a conceptual truth, the *analysans* has a *substantive* content. (The term 'substantive' is Etchemendy's; his paradigm cases of *substantive generalizations* are sentences like the former 'for every x, if x is a U.S. senator then x has a sister'.)

Before addressing Etchemendy's specific arguments, consider first how the discussion in the former section allows us to dispose of an unsophisticated version of it. Someone could argue in this way. "Although *that σ is logically true* is a conceptual truth, *that σ is true in every model* is not. For a model is a kind of set, and pure logic does not require that there exist any sets at all; hence, since the nonexistence of models is logically possible, it is logically possible that σ be (vacuously) true in every model, even though σ is obviously not logically true."[15] In other words: it is a *substantive*, hence *nonlogical* fact that there exists a model appropriate to refute the logical truth of σ. That is why the model-theoretic account is not satisfactory as an analysis, even though it could now and then get the correct extension of the properties it purports to analyze.

This argument is not good, because it assumes without proof that the existence of models (in particular, of the model needed to refute the claim that σ is a logical truth) is a nonlogical issue. But the model-theoretic account, as I have stressed in the former section, contends precisely the opposite of this. According to the model-theoretic account, the intuitive assertion that σ (a first-order translatable sentence) is not a logical truth is equivalent to the assertion that σ is not a truth in virtue of the meaning of the first-order logical constants in σ; and, the first-order semantics for the logical constants being a good semantic theory for their natural language counterparts, this contention already involves the claim that there exists a *preformal model* in which σ would not be true. The opponent of the model-theoretic account tries to seduce us into thinking otherwise (or perhaps deceives himself) by talking about *sets*, for there is a good argument to the contention that set-theory is much more *substantive* than pure logic. But what is here in dispute is whether or not the intuitive claim that a sentence is not a logical truth involves the existence of pretheoretical set-like entities, entities that admit of precise characterization with the help of set-theoretic apparatus. The advocate of the model-theoretic account believes that this is really so. In not addressing that issue, in hiding it behind the words "but a model is a kind of set," the argument above begs the question.

Not to beg the question, on the other hand, will prove to be tough. What must be established by the opponent of the account is that when we claim that, for instance, 'some U.S. senators are lawyers, some lawyers enjoy Jane Austen's novels, therefore some U.S. senators enjoy Jane Austen's novels' is not (intuitively) a logically valid argument, we are not claiming that there is a possible (set-like) domain, and possible (set-like) extensions in it of three different predicates, 'U.S. senator', 'lawyer', 'enjoys Jane Austen novels',[16] such that the premises are true and the conclusion false. For that is what the contention that the conclusion does not follow from the premises in virtue of the meaning of 'some' comes to, according to the defender of the model-theoretic account. And this does not seem so implausible on the face of it.

Now Etchemendy's argument is no more than a sophisticated version of the one just discussed. He asks us to consider any of the sentences saying, in the pure first-order language, that there are fewer than n objects in the universe, for arbitrary n. In these sentences there is no nonlogical expression (if we count the identity sign among the logical expressions) other than the "expression" that indicates the intended domain of quantification. Hence, the sentence that says that there are fewer than n objects would be a logical truth if the corresponding universally quantified sentence, "for every domain, there are fewer than n objects," were simply true, for arbitrary n. It is not; but only, Etchemendy claims, because we are making *extralogical* or *substantive* assumptions in our metatheory, asserting (by means of the axiom of infinity) that there are infinite domains.

To make his point clearer, Etchemendy uses this consideration: a finitist who adopted the model-theoretic account should have to say that, from a certain *n* onward, the sentence saying that there are less than *n* objects is a logical truth. But he need not. He could consistently keep his philosophy of mathematics while agreeing with us on logic. This purportedly shows that the account, taken by itself, overgenerates.

The sophistication Etchemendy introduces with respect to the former argument consists in considering a more doubtful case, that of the sentences whose nonlogicality requires, according to the account, infinite models, thus putting stress on the notion of *preformal model*. But the point is the same as before. Whether or not a sentence is logically true is a nonsubstantive matter. The model-theoretic account converts it into a substantive one. This is shown by means of the recourse to the finitist: the finitist and I do not disagree about questions of logic, but we do disagree about the existence of infinite models. The account, then, is wrong.

Now I do not agree that the discrepancy with the finitist is not logical. It is assumed that we agree that none of the sentences under issue is an intuitive logical truth. This is to say, according to the model-theoretic account, that none is true in virtue of the counterparts in them of the first-order logical constants. But *to the extent that the semantics we presuppose for them is correct*, that already involves the existence of an infinite preformal model.[17] Because if there is no infinite preformal model, it *does* follow from the meaning of the (counterparts of the) logical constants that some of these sentences are true. Thus the finitist must disagree with our semantics. And it is far from clear that *this* is not a logical disagreement. When defenders of finitism actually provide an alternative semantics for quantifiers, it does involve logical disagreement. On the other hand, the reasonable finitist cannot simply refuse entering the discussion (because, say, of the use we make of the terms in dispute) without explaining in some other way why he thinks that none of the sentences discussed is a logical truth. Once again, the defender of the model-theoretic account claims that, when the finitist does this, it will be clear that their disagreement concerns logical matters. The finitist cannot just say: "they are not logical truths, according to my intuitions, and according to these same intuitions the former claim has got nothing to do with sets, or set-like entities." This is not a question of intuitions. We have provided an explanatory theory for his, and our, logical intuitions. That can only be challenged with intuitions incompatible with the theory *and* supported by an alternative theory, at least equally explanatory; or with theoretical reasons to believe the theory false.

Therefore Etchemendy begs the question, as much as the proponent of the unsophisticated version of his argument did. The difference lies in the way they try to entice us. Etchemendy uses the apparent fact that the discrepancy with the finitist is a "substantive" one, where the former had recourse to the models being sets, and thus to the contentfulness of set theory, on that account. But it is a far cry from the substantiveness of facts about the composition of the U.S. senate to that of the facts disputed by finitism. As a matter of fact, the model-theoretic account involves the claim that logical issues are more substantive than they could seem at first sight. Merely to assert that they are not by putting forward examples like Etchemendy's finitist is to beg the question.

To sum up, to rely on the axiom of infinity is not, on my account to rely on an extralogical fact at all. The model-theoretic account (as the Frege-Russell account) does not rely on what domains of quantification consisting of actual entities there are. It relies on what domains of quantification are available, on what domains it makes sense to quantify over, because it wants to explain logical truth as truth in virtue of the meaning of the logical constants. If we quantify over natural numbers, or over vocabularies with infinitely many words, then those are possible logical values for the "expression" referring to the domain in any first-order argument; at least, they are so to the extent that the semantic analysis embodied in the account is a correct one.

My insistence that the model-theoretic account involves some modal element (the mild one involved in the contention that models are *possible logical values*) could strike an old chord in the reader. Am I not propounding, after all, a *representational* view of models? Let me try to answer this possible worry by explaining better than I did before exactly why the representational view is the wrong way to understand the model-theoretic account, now that I have explained how I understand that account. In the representational view of them, models represent possible logical values the nonlogical expressions could have, *given the meanings they actually have*. The problem with the model-theoretic account, viewing models this way, is not only, nor even mainly, that we are resorting to an unexplained notion of possible world, that we explain something we do not understand in terms of something we understand even less. The problem is that the model-theoretic account, on this view, is going to produce the mixing of phenomena that should be kept apart and should be given different explanations. Presumably, there is no representationally possible world in which 'this patch is entirely red and this (same) patch is entirely green' is true, but that has nothing to do with the reasons why there is no representationally possible world in which 'this patch is red or this (same) patch is not red' is true either.

According to my interpretation, in contrast, the actual meaning of the nonlogical constants does not matter at all, nor does it matter what logical values those expressions could have in some other possible worlds, keeping their meanings fixed. I do not use the idea of *possible world* at all. My possibilities are merely combinatorial possibilities, determined (to the extent that the consequence relation is determined) by the semantics of the language under study. What matters is the possible logical values of expressions with the logical syntax the nonlogical expressions have. (And we have an explanation why it matters, in our semantic account for the logical constants.) This is a purely semantic (or, if you prefer, *conceptual*) notion. In other words, nobody can (tacitly) understand the logical expressions without a (tacit) understanding of them.

Now it could seem that, although the notion of a possible set of logical values (i.e., truth values) for a given set of atomic sentences and that of a possible set of logical values (i.e., classes or any other set-like entity) for a given set of basic predicates, do not smuggle in any esoteric modality, everything is different with the notion of a possible domain of quantification. I would like to know an argument to this effect; I have not been able to find one in Etchemendy's book, nor can I think of one. To understand quantification in natural language, you must have the general idea of a possible domain of quantification, because, obviously, we do not quantify over "everything." There are many problems standing in the way of a correct account of this notion. It seems that the domain need not be the extension of a predicate. Perhaps it is equally incorrect to characterize it as a set. Or perhaps our linguistic practices leave us some leeway here. But I do not see how to obtain from this the result that the modality involved in that notion is more dangerous than the one we make use of when we speak of a possible combination of truth values for the atomic sentences in a propositional logic argument.

Let us assume that atomic sentences are given their truth values, either truth or false, by "atomic facts," "truth-makers" for atomic sentences. Let us assume further that there are only finitely many of them, say n, it being understood that this has the consequence that at most n atomic sentences can actually have a truth value. Neither assumption affects whether or not a sentence built out of $n + m$ atomic sentences, $m \geq 1$, 'not' and 'and' is a logical truth, according to my interpretation of the model-theoretic account. Suppose that the sentence is the negation of the conjunction of the $n + m$ atomic sentences. Then it is not a logical truth, because it is a conceptual fact strongly connected with the semantics of 'and' and 'not' that $n + m$ different atomic sentences could all be true. This is, I believe, as it should be.[18]

Etchemendy uses another argument to make his point. This one concerns the set-theoretic assumptions in Second-Order Logic. As is well known, taking the model-theoretic account relative to the standard semantics, there are sentences of second-order (formal) logic that would be logical truths if and only if substantive and disputed propositions of set-theory were true (like the Continuum Hypothesis, for instance). Etchemendy claims that here is where his contention is more obvious: the model-theoretic account transforms contentless, logical truths into substantive assertions.

This is again misguided. As I said before, to the extent that the analytical accuracy of the model-theoretic account is in question, logical theories for artificial languages are of interest only when they can be viewed as models (now using the word in the sense given to it by scientists, meaning "mock-up") of natural language, or of fragments of natural language. Set theory, as an adequate device for the building of logical theories, comes after the semantic theory of the language whose logic we are investigating (perhaps just the same language in which we develop set theory itself). Set theory, by itself, cannot determine what models there are, now using 'model' in the sense previously defined in this paper. The semantics of the language reveals what is a possible domain of quantification, what a possible extension for the sentential functions, etc. The semantics of the language, to sum up, determines the clear cases in the extension of the relation of logical consequence. Set theory is only a tool to give us a more precisely defined sense of 'model'; a tool to build, in a precise way, mock-ups of natural language and of what I have called *preformal models*. As I said before, it is foreseeable that the semantic properties of the language under study do not determine everything about the logical consequence relation (because the meaning of any expression is more or less vague). Thus in constructing a precise logical theory for a language, we can make decisions not determined by its semantics; we can, to a certain extent, create a new language, or reform our natural one.

To this extent the logical theory is not only determined by, but actually shapes, our logical intuitions (and their semantic underpinnings). But, leaving aside this leeway, set theory has nothing to say about the extension of the consequence relation.

Therefore, to decide whether the second-order case can be used to mount an attack on the model-theoretic account, we must first of all find good counterparts of sentences in natural language for the problematic, second-order sentences. The most important thing is to find good counterparts of the specific second-order logical constants, and to decide that the formal second-order semantics constitutes a good semantic theory for them. This is not easily done. Of course, the fact that in certain arguments we quantify over properties or sets is no sign of "second-orderness." Quantification over abstract entities like numbers or properties is, in principle at least, first-order quantification. There is no semantical difference between the reasons that make 'John loves every woman, Nancy is a woman, so John loves Nancy' a valid argument and the reasons that make 'Napoleon has all properties of a good general, being brave is a property of a good general, so Napoleon is brave' a valid argument. At least, there is no semantical difference that we can adequately capture representing the first as a first-order argument and the second as a second-order one. What we must do is to find, in natural language, expressions that seem to work as quantifiers over arbitrary subsets of the intended domain of objects. Now there are good arguments to the effect that plurals, and particularly plural quantifiers, are such expressions.[19]

Once we have settled that, there are several courses we can follow in view of facts like the ones mentioned by Etchemendy. One is to claim that second-order logic is not logic. We could claim that in the case of second-order logic we have gone too far away from what we should be willing strictly to call "logic." Inspiration along those lines comes in the arguments offered by Jané [9]. Another possible course is to bite the bullet, endorsing arguments like the ones given by Shapiro [11] against the substantivelessness of logic. The argument in this paper drives me toward this last position. But the important point is that, whatever the course chosen, Etchemendy's criticism does not apply without further ado.

5 To conclude, I will discuss briefly the historical question. Did Tarski maintain Etchemendy's version of the interpretational account, or had he in mind something more like the interpretation that I have proposed? The issue is not clear to me. While reading "On the Concept of Logical Consequence" (and certain passages of "The Concept of Truth in Formalized Languages"), I find myself uncertain how to interpret his proposals. On the one hand, Etchemendy seems right in that he is giving a version of the substitutional account, modified by means of the semantic devices he himself discovered in order to avoid the problem of overgeneration due to the lack of expressive power in the language. On the other hand, we find passages like this:

> Certain considerations of an intuitive nature will form our starting point. Consider any class K of sentences and a sentence X which follows from the sentences of this class. From an intuitive standpoint it can never happen that both the class K consists only of true sentences and the sentence X is false. Moreover,

> since we are concerned here with the concept of logical, i.e. *formal*, consequence, and thus with a relation which is to be uniquely determined by the form of the sentences between which it holds, this relation cannot be influenced in any way by empirical knowledge, and in particular by knowledge of the objects to which the sentence X or the sentences in the class K refer. The consequence relation cannot be affected by replacing the designations of the objects referred to in these sentences by the designations of any other objects. The two circumstances just indicated, which seem to be very characteristic and essential for the proper concept of consequence, may be jointly expressed in the following statement: (. . .) ([13], pp. 415–415.)

Then he gives the substitutional account, which he subsequently turns into his own semantical account in order to avoid depending on the expressive power of the language for which the definition is given.

As I said, I do not know exactly what to make of the quoted paragraph. What is true is that, for Tarski, 'formal' does not mean 'purely syntactic'. He usually refers to purely syntactic characterizations of whatever property as 'structural'. Moreover, one of his explicit aims in both papers, "The Concept of Truth in Formalized Languages" and "On the Concept of Logical Consequence," is to argue against syntactical characterizations of the concepts of truth and of logical consequence. What then did he mean by 'formal'?

I stressed before that in the proposal I have made the logical consequence relation, in opposition to the more general analytic consequence relation, has indeed a "formal" character. That formal character derives from the "contentless" meaning of the logical expressions, according to the meaning that a correct semantic theory attributes to them. Now, the main part of a correct account of the logical properties, as I see it, is a good theory that gives the meaning of the logical constants. Well, *Tarski's definition of truth for a first-order language embodies such a theory*. In Frege's or in Russell's work we can find good *elucidations* of the semantics of the logical particles. But it is only in Tarski's that we encounter a good *theory*. Or so I believe, and here there is another point of discrepancy between Etchemendy and me.[20]

I am not claiming that my explanation of why logical truths are formal corresponds neatly to what Tarski had in mind. It is difficult to claim something like that, in part because probably only some more or less rough surmise buttressed Tarski's contention about the "formality" of the intuitive consequence relation. But I think that his contention, together with the fact that we can give a clear explanation of it, compels us to have a more sympathetic look to Tarski's findings than Etchemendy allows. In Chapter 6 of [6] (and earlier in [4]) Etchemendy criticizes what he calls 'Tarski's fallacy'. The fallacy is a harsh modal one, and it is present, according to Etchemendy, in a passage of "On the Concept of Logical Consequence" in which Tarski is trying to show that his analysis captures the "modal knot" present in the intuitive concept of consequence. This is the passage. I quote it from Etchemendy's book, with Etchemendy's translation and emphasis:

> It seems to me that everyone who understands the content of [my] definition must admit that it agrees quite well with ordinary usage. This becomes still clearer from its various consequences. *In particular, it can be proved, on the basis of this definition, that every consequence of true sentences must be true,*

> and also that the consequence relation . . . is completely independent of the sense of the extralogical constants which occur in these sentences. ([6], p. 86.)

Etchemendy represents Tarski's inference here as a well-known fallacy, inferring from a sentence with the structure,

$$\Box(p \rightarrow (q \rightarrow r))$$

another with the structure,

$$p \rightarrow \Box(q \rightarrow r).$$

Let us say that an argument Γ has *Tarski's property* if and only if, for every assignment of "meanings" to the nonlogical terms in Γ such that all the premises of Γ are true, the conclusion of Γ is also true. It is then indeed necessary that, if any argument, say, Γ, has Tarski's property, then, if the premises of Γ are true, then the conclusion of Γ is also true. It is necessary, for the simple reason that one of the possible assignments of meanings is the actual one; and so, of necessity, given that Γ has Tarski's property, if the premises (with their actual interpretation) are true, then the conclusion is surely true. But it does not follow from this that if Γ has Tarski's property, then, necessarily, if the premises of Γ are true, the conclusion of Γ is also true. That is, it does not follow from the foregoing obvious definitional fact the conclusion that *to have Tarski's property* suffices for (i.e., *explains*) the modal fact that the conclusion of a logically correct argument *must* be true, given that the premises are true.

If, instead of considering arguments, as Tarski does, we consider sentences, as we have been doing until now, it is still easier to see the fallacy that Etchemendy attributes to Tarski, and to see that it is a fallacy. It is necessary that, if any sentence, say, σ, has Tarski's property (whatever the meaning of the nonlogical terms in it could be, it would be true), then the sentence σ is true. It is as necessary as a simple definitional fact could be. But it does not follow from that that, if sentence σ has Tarski's property, then sentence σ is necessarily (logically) true. It does not follow that *to have Tarski's property* suffices for (i.e., *explains*) the fact that the truth of sentence σ is a necessary (logical) one. This is, Etchemendy argues, to infer a sentence with the form $\ulcorner(p \rightarrow \Box q)\urcorner$ from another with the structure $\ulcorner\Box(p \rightarrow q)\urcorner$. And that is, indeed, a modal fallacy.

But it is more than a modal fallacy. To put it in those terms makes it less bizarre than what it in fact is. Put in those terms, it seems the kind of fallacy that even the best logician could commit. With my comments, though, I have tried to point out that it is much worse than that. It is to infer the correctness of an analysis simply from a definitional fact, from a sheer stipulation. And I think that to charge Tarski with committing that kind of fallacy is to apply too little charity.

Now it is obvious from the context that, in the last quoted text, Tarski is referring to the two intuitive properties that the ordinary concept of consequence has which he himself refers to in the text quoted before. Those properties were, first, what I have called the "modal knot" and, second, the alleged "formality" of the consequence relation. Etchemendy seems to think that when Tarski says that "it can be proved" that the defined consequence relation has both intuitive properties, he is speaking about a rigorous kind of proof. Indeed, the only rigorous proof which one could come up with is the fallacious one he describes and I have outlined above. But I think that the fact that the "it can be proved" also covers the second intuitive property that the definition allegedly captures, namely, the 'formal' character of the intuitive consequence relation, allows us to disregard Etchemendy's interpretation. Only in a vague sense of 'proof' could it be proved that Tarski's definition captures that 'formality', because the relevant sense of 'formality' has not been made explicit. Thus it is only in a similar, ordinary, nonrigorous way that, according to Tarski, it could be proved that his definition explains the "modal knot" of the intuitive consequence relation.

And what kind of proof could that be? My interpretation of the model-theoretic account furnishes one, the one I gave above. Because of that, I think it is better to believe that Tarski had in mind and was pointing to a "proof" like the one I have espoused. If possible, it is always better to interpret influential thinkers in a way that contributes to understanding their influence.[21]

Acknowledgments Among the many people who have helped me in articulating the ideas in this paper, I would like to express my gratitude to George Boolos, Richard Cartwright, John Etchemendy, Ignacio Jané, Begoña Navarrete, Jamie Rucker, and Robert Stalnaker. The paper has greatly improved both in truthfulness and in clarity because of their comments, criticisms and suggestions; any remaining falsehood or confusion, however, should most probably be blamed on my stubbornness. The first version of the paper was completed while I was a visiting scholar to the Center for the Study of Language and Information, Stanford University. I thank the CSLI for the warm and encouraging atmosphere I found there, and the Spanish DGICYT for the grant that made possible my stay at the CSLI. This paper is part of the research project PB90-0701-C03-03, funded by the DGICYT.

NOTES

1. It is well known, and it is frequently mentioned, that Frege was from the very beginning of his work well aware of the necessity of distinguishing rules of inference from sentences in any logical system, whereas Russell only in an openly painstaking way came to realize it. What it is not so frequently acknowledged is that this does not place Frege in a better position regarding the problem discussed in the main text, namely, the asymmetry in the explanatory adequacy of his account of the validity of valid arguments versus that of valid propositions. I think that, on the contrary, Russell's pains were partially due to his realizing the difficulty, whereas Frege does not even seem to have been aware of it. Frege's *logical* rigor is frequently opposed to Russell's sloppiness. It is only fair to say that, in matters of philosophy, the roles are *quite* inverted.

2. For historical nuances, see Hodges [8].

3. Typically L will be a "formal language," an artificial creation. In many mathematically interesting applications, L will have little similarity to any fragment of natural language. But in the context of this discussion, we are interested only in formal languages that are intended as *models*—in the scientist' sense of the word—or "mock-ups" of natural language. This is so because we are discussing to what extent the model-theoretic account is a good analysis of the intuitive concept of consequence, and the latter applies primarily to arguments in natural language.

4. See Chapter 5 of [6] for a careful discussion of the idea of *different interpretations* applied to quantifiers. Intuitively, quantifiers are "logical expressions," so they must be counted among the ones in F, the ones whose interpretation we do not change. But we need to take account of different universes to say, for instance, that none of the sentences saying that there are more than *n* objects (for each *n*) is a logical truth. (I discussed this problem before, when commenting on Wittgenstein's and Ramsey's criticism of the Frege-Russell account.) And to run the model-theoretic account, as Etchemendy presents it, we need to find some expression which is not in F and the "meaning" of which is the intended domain of quantification. That is where the idea of the universe as the interpretation of quantifiers comes from. In English, we could say that the 'thing' part of most quantifier expressions has as its interpretation the domain, and is not in F, whereas the other part remains among the expressions in F. There are other natural languages that do not provide for such a trick. But this is not a problem either. After all, there is nothing strange in the fact that in the logical form corresponding to an utterance there is some symbol which does not correspond to anything in the sentence uttered. In the logical form of the utterances made by means of many road signs there must be a singular term (meaning *here*, or *here and now*) to which nothing corresponds in the sign itself.

5. I assume the usual objectual truth-conditions for quantified sentences, with an appropriate extension of the satisfaction definition to second-order quantified sentences. Some of the variables may be second-order.

6. See Chapter 7 of [6].

7. Tarski is no more clear than I have been about what he meant by the "formality" of the intuitive concept of consequence. The issue will play an important role in my discussion of the historical question in the last section of the paper.

8. I am not myself endorsing this view, but only stating Etchemendy's. As the next section will make clear, to me this is not a real weakness in the model-theoretic account, but only a symptom that Etchemendy—and perhaps other people with him, even Tarski maybe—is not correctly stating the interpretational view of the model-theoretic account. From the understanding I provide in the next section, cross-term restrictions on possible interpretations follow smoothly.

9. In the last two paragraphs, I have tried to condense what Etchemendy [6] says in Chapter 7, pp. 100–106, and in Chapter 8.

10. There is no mystery in assuming (as I do) that there is no conceptual difficulty in the idea of a theory of meaning.

11. To infer soundly *according to the semantic rules of the language*. There are other senses of *sound inference* that we must carefully distinguish from the former; for instance, *to infer soundly according to the rules of a certain game we are playing* or *according to the laws of nature*, etc. A semantic theory has little to do with the explanation of the validity of these other kinds of inference.

12. Due to this "universality" of the logical constants, and due to the fact that in building a logical theory we abstract from irrelevant particularities of the language under study, what we say about the logical properties of an English fragment applies immediately to almost any other fragment of natural language. We could as well talk about the logical properties of "pure thoughts," if by that we did not intend anything about thoughts mystically separated from their linguistic dresses, but only that our theories apply widely across a whole range of languages. Also in this regard, I must repeat here a point I made before, to assuage the mathematical logician who,

happily working miles away from the vagaries and idiosyncrasies of natural language, listens with understandable reluctance to any worry about the relationship between formal and natural languages. What we are discussing here is the extent to which the model-theoretic account constitutes a good, precise formalization of some informal concepts. As the preformal concepts apply to arguments in natural language to which the formal ones are more or less rough counterparts, what we need to understand is what explanation can be drawn from the formal account of the basis for the application of those preformal concepts.

13. If the semantic hypotheses about the logical constants in natural language we in effect make when we build first-order logic, or when we say that first-order logic captures the consequence relation restricted to a certain fragment of natural language, are correct, then we have a more precise explanation of what we meant when speaking about the "contentless" character of the logical expressions. Their meaning, their contribution to the truth conditions of every sentence in which they appear, is made only in terms of very abstract semantic features of the "contentful" expressions. They are "sensitive" only to truth values or to extensions in an intended domain.

14. Although, as it is clear from the paper, I align myself with the supporters of the semantic view of the logical properties, I believe that the building of deductive calculi is also an integral part of logic. Even if σ follows from the sentences in Γ analytically, we can distinguish two cases: the case in which it is obvious that it is so, and the case in which it is not obvious. This is not a sheer psychological distinction, although it undoubtedly has anthropomorphic elements. It could be that the step from Γ to σ requires merely taking into consideration the semantic rule for a term, or it could be that it requires many different semantic rules, applied in a convoluted way. It is similar to the difference in being able to "see" a checkmate when it requires the obvious movement of a piece and when it requires several unexpected movements. I believe that the search for proof procedures, like the several "natural deduction" calculi, has to do with this "reducing to minimal semantic complexity." Those calculi allow us to pass from Γ to σ in such a way that each step is semantically obvious. This has nothing to do, on the other hand, with the idea that logical truth and logical consequence are purely syntactic concepts, as some people used to believe, or with the equally misguided idea that those properties are to be explained in terms of verification procedures, as some other people believe nowadays. What is wrong with these views is the belief that the justification for the validity of an argument lies in some abstract "rule of inference." The justification actually lies in the semantics of some of the expressions already present in the argument.

15. A similar argument can be found in McGee [10].

16. Bear in mind that it does not matter that the extensions in question *are really constituted of* U.S. senators, lawyers, and enjoyers of Jane Austen's novels. The model-theoretic account does not have us contending, in claiming that the argument is invalid, that there exists a *possible world* in which some U.S. senators are lawyers, some lawyers enjoy Jane Austen's novels, but no U.S. senator enjoys Jane Austen's novels. That may be so, but it has nothing to do, according to the account, with the invalidity of the argument. (And in other cases of logically invalid arguments it is not sensible to say that the relevant *possible world* exists.) The pursuit of *representational semantics*, as Etchemendy stresses, is a red herring in this context.

17. Well, that need not be true. But other sentences would allow us to make the point.

18. In personal conversation, Etchemendy objected that, according to my interpretation, the model-theoretic account would not explain the "modal knot" in the intuitive concept of consequence. The problem would lie in the use I make of the modal notions of *possible domain* and *possible logical value*; it seemed to him that I was stretching the idea of *semantically (or conceptually) possible* too far in saying that the possibility of quantifying over infinite domains is authorized by semantics. The previous three paragraphs are intended to answer this criticism. The point is that although the model-theoretic account does involve modal claims, they are not modal facts about what worlds are possible (as they are when we take the models representationally). They are just facts about what combinations of truth values for atomic sentences, or what possible domains of quantification are semantically allowed. These facts may well be *substantive*, but their substantiveness does not stand in the way of offering an understanding of the altogether mysterious "modal knot" in the intuitive concept of consequence.

19. See Boolos [2]. There is something I find rather difficult to swallow in Boolos' proposals, though. He wants the expressive power of second-order logic, especially to formalize set theory. On the other hand, he does not find appropriate the set-theoretic semantics in this particular case, and he gives good reasons for this. Then, having showed that plural quantifiers are good candidates as counterparts in natural language of second-order quantifiers, he uses them, with their intuitive semantics, to give a different semantics for second-order *formal* logic. (He does this by means of a translation procedure, from formal second-order logic to natural language, in the aforementioned paper, and by means of an informal truth theory, using plural quantifiers in the metalanguage in [3].) I do not find this adequate. The informal understanding of plural quantification, I believe, does not take us far enough to interpret the sentences of second-order logic. We would need a precise semantic theory for them. (We have a good intuitive understanding of simple sentences involving first-order quantifiers, but our intuitions blur or simply disappear as soon as we consider complicated sentences. Nevertheless, we can judge that the first-order semantics—or something very much like it, say, the "Generalized Quantifiers semantics" of Barwise and Cooper [1]—correctly captures the semantics of quantifiers in the simple sentences we can judge about; then, once we have the explicit, systematic theory, using it we get a clear understanding of whatever sentence we confront. This is what we would need for plural quantifiers.) And the best one I can think of is the usual semantics for formal, second-order logic. It is by using it that I am able to parse moderately complicated sentences with plural quantification. I grant that we should make an exception when our intended domain is not a set. And I do not have clear ideas about what kind of exception this is, and what the consequences are of its being an exception.

20. In [5] Etchemendy argues that Tarski's definition of truth does not have any semantic import. I believe that this is also mistaken, but I cannot argue against it here.

21. Sher makes a similar criticism of Etchemendy in Chaper 3 of [12].

REFERENCES

[1] Barwise, J. and R. Cooper, "Generalized quantifiers in natural language," *Linguistics and Philosophy*, vol. 4 (1981), pp. 159–218.

[2] Boolos, G., "To Be is to be the value of a variable . . . or some values of some variables," *Journal of Philosophy*, vol. 81 (1984), pp. 430–449.

[3] Boolos, G., "Nominalistic Platonism," *Philosophical Review*, vol. 94 (1985), pp. 327–344.

[4] Etchemendy, J., "The doctrine of logic as form," *Linguistics and Philosophy*, vol. 6 (1983), pp. 319–334.

[5] Etchemendy, J., "Tarski on truth and logical consequence," *The Journal of Symbolic Logic*, vol. 53 (1988), pp. 51–79.

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

[7] Hacking, I., "What is logic?," *Journal of Philosophy*, vol. 76 (1979), pp. 285–319.

[8] Hodges, W., "Truth in a structure," *Proceedings of the Aristotelian Society*, vol. 86 (1985), pp. 135–151.

[9] Jané, I., "A critical appraisal of second-order logic," forthcoming in *History and Philosophy of Logic*.

[10] McGee, V., "Two problems with Tarski's theory of consequence," preprint.

[11] Shapiro, S., *Foundations without Foundationalism*, Oxford, Clarendon Press, 1991.

[12] Sher, G., *The Bounds of Logic. A Generalized Viewpoint*, Cambridge, MIT Press, 1991.

[13] Tarski, A., *Logic, Semantics, Metamathematics*, Oxford, Clarendon Press, 1956.

[14] Wittgenstein, L., *Tractatus Logico-Philosophicus*, English translation by D. F. Pears and B. F. McGuinness, Ithaca, Cornell University Press, 1971.

Departamento de Lógica, Historia y Filosofía de la Ciencia
Universidad de Barcelona
Baldiri Reixach, s/n 08028 Barcelona (Spain)