Are There Model-Theoretic Logical Truths that Are not Logically True?*

MARIO GÓMEZ-TORRENTE
Instituto de Investigaciones Filosóficas, UNAM, México DF 04510, Mexico
mariogt@servidor.unam.mx

The question in the title is similar to other important unsettled questions prompted by attempted mathematical characterizations of pretheoretical notions, by coextensionality "theses" about them.[1] Church and Turing's thesis that a function is computable iff it is recursive[2] gives rise to the question "are there computable functions that are not recursive?". The thesis, especially associated with Stephen Cook, that a natural problem has a feasible algorithm iff it has a polynomial-time algorithm[3] gives rise to the question "are there natural problems having a feasible algorithm that do not have a polynomial-time algorithm?".[4] But there are remarkable dissimilarities too. A central one is that the notion of "model-theoretic logical truth" is a notion relative to (at least) a choice of a set of formalized languages, of a set of logical constants, of a notion of model, and of a notion of truth in a model, while the notions of recursiveness and of a polynomial-time algorithm are

* Parts of this paper were presented in workshops at the University of California at Irvine (2002), the Universidade Nova de Lisboa (2003), the University of Melbourne (2004), and the Universidad de Santiago de Compostela (2006), and as lectures at the Universidad Nacional Autónoma de México (2006) and the Universidade Federal do Rio de Janeiro (2006). I thank the audiences at these events for very helpful comments and criticism. Special thanks to Rodrigo Bacellar, Bill Hanson, Øystein Linnebo, Gila Sher and two anonymous readers.

[1] Here by "pretheoretical" of course I don't mean "previous to any theoretical activity"; in this sense there could hardly be pretheoretical notions of computability, of a feasible algorithm, or of logical truth. What I mean is "previous to the theoretical activity of mathematical characterization".

[2] See Church (1936) and Turing (1936/7). Of course, the notion of recursiveness is provably coextensional with the notions of lambda-conversion and Turing-machine computability used respectively in the enunciation of the original theses of Church and Turing.

[3] See especially Cook (1991); see also Cook (1971).

[4] The three converse questions have some interest too, but it is widely conceded that properly understood they must have a negative answer. In the case of logical truth, it is widely conceded that, provided one constructs one's model theory with sufficient care, there cannot be pretheoretical logical truths that are false in some model.

not relative to anything in any such conspicuous way. Nevertheless, arguments against particular coextensionality theses, involving fully relativized notions of model-theoretic logical truth, will be significant provided the theses are *prima facie* reasonable and/or have been actually proposed by logicians.[5]

One common way of arguing against coextensionality theses involving fully relativized notions of model-theoretic logical truth has been by arguing that some model-theoretic logical truths in the relevant sense are not necessary or are not *a priori*, for under most pretheoretical conceptions of logical truth a logical truth must be necessary and *a priori*. One important fact in the background of this paper is that the process of formalization leaves unclear, though perhaps not undetermined, the answers to some questions one must answer before one can ask if an interpreted sentence of a formalized language is necessary or *a priori*. This is not in itself a weakness of formalization, since the main aim of formalization is to obtain formal sentences which, unlike their correlates in natural language, have a truth-conditional content made absolutely precise by stipulation; in this process, features of the natural language sentences relevant to their modal character and their epistemology are simply abstracted from. But then formalization does have the result that some questions about the modal and epistemic status of a model-theoretic account of logical truth do not have a clear, or perhaps even a determinate answer. One aim of this paper is to emphasize this often neglected fact. The paper's main aim, however, is to argue that, once one lays open some natural or at least plausible ideas about the modal character and the epistemology of the classical first-order quantifiers, some *prima facie* reasonable coextensionality theses are false or at least must be somewhat qualified.

[5] I share the view, especially associated with Kreisel (1967), that one can give informal but potentially conclusive arguments both against *and* for theses asserting the coextensionality of pretheoretical and theoretical concepts. In this paper I will be especially concerned with giving tentative arguments *against* certain particular coextensionality theses in the case of model-theoretic logical truth. (Kreisel's own argument *for* one of these coextensionality theses will be mentioned below.) At least until recently, an unKreiselian view has been widespread in the case of the Church-Turing thesis. The view seems to have been widely held that while this thesis can be refuted, it cannot be conclusively argued for, since it relates a theoretical concept and a pretheoretical one that cannot be used in rigorous general reasonings. But a Kreiselian view of the Church-Turing thesis has been urged by Harvey Friedman, Saul Kripke and others.

In particular I will argue that, given those ideas, the specific coextensionality thesis put forward by Tarski, the main proponent of the model-theoretic method for the mathematical characterization of logical truth, doesn't hold even though it is *prima facie* reasonable. In order to get to this conclusion, I will first enunciate and distinguish a number of coextensionality theses that sound Tarskian somehow, and I will offer a quick evaluation of each (in section I). For each of these theses I will claim that either it is weaker or stronger than Tarski's thesis. In the course of this examination of theses I will survey some previous critiques of model-theoretic characterizations of logical truth, and will find them unsatisfactory. In section II I will state the thesis that most deserves the name 'Tarski's thesis', and I will note that, under natural assumptions, there are model-theoretic logical truths in the sense relevant to Tarski's thesis that are not necessary, and hence not logically true under most conceptions of logical truth. Some of those natural assumptions include assumptions about the modal behavior of the classical first-order quantifiers, that are not part of their explicit classical extensional model theory. But, as advanced above, some such assumptions need to be made explicit before one can meaningfully ask the question whether a classical quantificational sentence is necessary. And I will argue that the assumptions I will use are the most natural given the classical extensional model theory.

Similarly, one needs assumptions about the epistemology of classical first-order quantifiers before one can ask the question whether a classical quantificational sentence is *a priori*. The final section III describes some assumptions about the epistemology of classical first-order quantifiers that, though plausible, are potentially more controversial than the assumptions about their modal behavior described in section II. I note that, under these potentially controversial assumptions, some model-theoretic logical truths in the sense relevant to Tarski's thesis, and even in the sense relevant to some weaker theses described in section I, are not *a priori*, and hence presumably not logically true. Nevertheless, it appears that those theses need only be slightly qualified in order to free them from the counterexamples I will offer.

I

An especially important coextensionality thesis that Tarski held, but that is clearly not his (strongest) coextensionality thesis, is the following:

(T1) A sentence of a classical propositional/quantificational language is logically true in the pretheoretical sense iff it is true in all classical propositional/quantificational models which (re)interpret its constants (other than its classical propositional/quantificational logical constants).

(T1) is very specific about the class of sentences it talks about: the sentences of classical propositional and quantificational languages, both first- and higher-order. It is also very specific about the class of models it talks about, and about the notion of truth in a model that is at stake, which are just the classical, Tarskian ones[6]: in particular, a propositional model is any assignment of values from the set {Truth, Falsehood} to the propositional letters, and a quantificational model is seen as a sequence composed of a set-domain of quantification built out of existing objects, plus extensions drawn out of this domain for the predicate, function, and individual constant letters of a language in the relevant class. Also, (T1) talks about an absolutely specific set of logical constants, namely: the truth-functional propositional connectives and the classical quantifiers of finite order (plus the predicate of identity and/or a predicate of intra-typical membership in some formulations).

Is (T1) true, or is it false? Most people until relatively recently have thought that it must be true (or at least most of those who agree that the higher-order quantifiers are logical constants have thought that it must be true). Both critics and defenders of the model-theoretic approach (including myself in previous work) have tended to agree that variations of Kreisel's (1967) argument put beyond reasonable doubt (T1) as restricted to propositional and first-order logical constants. (See Etchemendy (1990), ch. 11, Hanson

[6] There are, of course, some doubts voiced in the literature about whether Tarski's 1936 notion of a quantificational model is what I'm calling 'the classical notion of a quantificational model'. (See, e.g., Etchemendy (1988), Bays (2001), Mancosu (2006). For alternative views see Gómez-Torrente (1996) and Ray (1996).) In my view there are in fact differences between Tarski's notion and the current notion (see e.g. Gómez-Torrente (2000)), but these are not the differences purportedly detected by the doubters. One of these differences is that in 1936 Tarski required, as a precondition for the applicability of his theory of logical consequence, that the domain of an interpretation of a first-order language be denoted by a non-logical predicate of the language, and this convention is not used with the current notion; I will come back to this convention. But it is relatively uncontroversial, at any rate, that at some point Tarski adopted all the now common conventions about models, and that (T1) (and (T1(1)) below) were Tarskian theses.

(1997), Gómez-Torrente (1998/9).) To be precise, the following is what has been thought to be beyond doubt:

> (T1(1)) A sentence of a classical propositional/first-order quantificational language is logically true in the pretheoretical sense iff it is true in all classical propositional/quantificational models which (re)interpret its constants (other than its classical propositional/first-order quantificational logical constants).

In particular, the question whether all model-theoretic logical truths in the sense of (T1(1)) are logically true has been thought to receive a positive answer by the following Kreiselian argument: let *S* be a propositional or first-order model-theoretic logical truth; then, by the completeness of propositional and first-order logic, *S* is derivable without premises in a wide array of deductive calculi; and for any of these calculi one can easily check by inspection that they can only yield sentences that strike one as logically true, under a wide variety of pretheoretical conceptions of logical truth. Nevertheless, in section III we will see a possible qualification to the conclusion of this argument.[7]

Recently several people have given arguments purporting to show that certain higher-order quantificational sentences are true in all classical quantificational models which (re)interpret their constants (other than their classical logical constants) and yet are not logically true. Etchemendy (1990), ch. 8, and McGee (1992) are perhaps the foremost examples of proponents of alleged counterexamples to (T1). The issue is a subtle one, but it seems fair to say that these attempted refutations have not gained anything close to a wide acceptance. McGee's alleged counterexample, in fact, is based on assumptions which go against the received view in set theory. My own view is that Etchemendy's and McGee's alleged counterexamples are unconvincing. (I give a critical discussion of these

[7] Another qualification is in any case needed in view of sentences such as '$(\exists x)(P(x)\vee{\sim}P(x))$', which are true in all (non-empty) models but may not be logically true. One can of course relax the convention of not contemplating empty models and prove suitable completeness theorems for the corresponding slightly non-standard calculi and appropriate variations in the notion of truth in a model. (See e.g. Quine (1954).) But one can also adopt a reasonable (yet apparently unexplored) view (described in a later note) according to which sentences such as '$(\exists x)(P(x)\vee{\sim}P(x))$' are not properly interpreted (or do not express propositions) when no non-empty domain for the quantifiers has been provided.

counterexamples in Gómez-Torrente (1998/9). See also Soames (1999), ch. 4, for specific discussion of Etchemendy.)

The alleged counterexamples and the general arguments supporting (T1) against them are too sophisticated for me to go into them here without digressing excessively. But what I want to claim for the moment is not that (T1) has not been refuted so far. I want to claim that (T1) is too weak to be called 'Tarski's thesis' (despite being so called both by McGee (1992) and by myself in Gómez-Torrente (1998/9)). (T1) is not the strongest coextensionality thesis Tarski postulated. The reason why it is too weak is that it talks about a very restricted set of logical constants. Tarski was clearly not concerned with the statement of a thesis about a set of logical constants so severely restricted by stipulation (of a list) as the set mentioned in (T1). To be sure, he contemplated the possibility that the notion of a logical constant might be so hopelessly obscure as to make arbitrary any delimitation of the borderline between logical and non-logical constants. But his later work (e.g., Tarski (1966)) shows that he never quite accepted that possibility. Tarski would have been ready to accept that other constants besides the classical logical constants of quantificational languages are logical constants, even assuming that the borderline between logical and non-logical constants may be fuzzy to some extent.

Tarski's thesis was something which, unlike (T1), is reasonably liberal about the class of logical constants it talks about. Let's consider this:

> (T2) A sentence of a formal language which possibly extends a classical propositional/quantificational language with new logical constants which are propositional connectives, quantifiers or predicates is logically true in the pretheoretical sense iff it is true in all classical propositional/quantificational models which (re)interpret its constants (other than its logical constants).

(T2) is just like (T1), but it does not restrict itself to any severely limited set of logical constants; consequently, it also does not restrict itself to sentences of classical quantificational languages, but talks about sentences with possibly new logical constants which are propositional connectives, quantifiers, and predicates having the same syntax as their analogues in classical quantificational languages. For example, one of the languages (T2) talks about is a typical quantificational modal language.

A decisive problem with (T2) is that, no matter how one understands the notion of truth in a model that appears in its formulation, and given a natural choice of logical constants, it is obviously false, and it is pretty absurd to think that Tarski might have had something like this in mind. Classical propositional or quantificational models are clearly not appropriate for a theory of the logical properties of non-extensional logical constants, such as '□' ("it is necessarily the case that"). To make the problem vivid, concentrate on a simple example. Think of a classical propositional language with 'p' as the only propositional letter (and the only non-logical constant), and add to it '□'. A classical propositional model for this language simply assigns a truth-value to 'p'. So there are just two classical models for this language: M_T which assigns Truth to 'p', and M_F which assigns Falsehood to 'p'. And thus there are just four possible combinations of truth-values in M_T and M_F for the formula '□p': (1) '□p' is true both in M_T and in M_F; (2) '□p' is true in M_T and false in M_F; (3) '□p' is false in M_T and true in M_F; (4) '□p' is false both in M_T and M_F. Given (1) or (2), the sentence '(p⊃□p)' is true in all classical models, and it is thus a counterexample to (T2), for it is not a pretheoretical logical truth on any reasonable conception of logical truth; given (3) or (4), the sentence '(p⊃~□p)' is true in all classical models, and for the same reason is again a counterexample to (T2).

So Tarski's thesis was something weaker than (T2). One possibility would be to refine (T2) by specifying a finite set of extensional and non-extensional constants we want to make our claim about and by trying to be specific about some corresponding suitable non-classical notions of model plus accompanying notions of truth in a model. For example, consider this:

> (T3) A sentence of a classical propositional/quantificational/modal language is logically true in the pretheoretical sense iff it is true in all propositional/quantificational/Kripke models which (re)interpret its constants (other than its classical propositional/quantificational/modal logical constants).

(T3) may sound reasonable and entrenched in logical practice. In this respect it seems to me to be very much like (T1). This is not to say that it has not been criticized in the literature. A prominent example of a criticism of this kind can be derived from Ed Zalta's (1988). Zalta gave a type of sentences which could be seen as counterexamples to (T3) if we accepted his considerations and used them from a certain point of view. (But Zalta did not construct his examples as counterexamples to the adequacy of the model-theoretic method.[8]) Here is perhaps the simplest of his examples (the others are essentially identical).

Think of a propositional modal language which includes a monadic sentential operator 'A' taken as a logical constant meaning "it is actually the case that". For such a language there is a somewhat standard Kripkean definition of model and of truth in a model, and hence also of model-theoretic logical truth as truth in all models. A model for such a language is a quadruple of the kind (W, R, α, V), where W is a set (intuitively, a set of worlds), R a binary relation on W (intuitively, the relation in which two worlds w_1 and w_2 stand when all propositions true in w_2 are possible in w_1), α a member of W (intuitively, the actual world of the model), and V an assignment of truth-values to pairs propositional letter-world. V can be extended to a full assignment of truth-values to every pair formula-world by means of certain well-known recursive satisfaction clauses for the logical constants of the propositional modal language. A sentence is called true in a model of this kind if it is assigned the value Truth in the world α of the model. And a sentence is called logically true in the model-theoretic sense if it is true in all models (or in all models where R verifies a certain property).

The recursive satisfaction clause for formulae of the form 'Aφ' says that they are assigned the value Truth in a world w of a model of this kind if φ is assigned the value Truth in the world α of the model. This means that every formula of the form of

(1) $Ap \supset p$

will be assigned the value Truth in the world α of any model, since if 'p' is false in α then 'Ap' is false in α. So every formula of the form of (1) is a model-theoretic logical truth given the somewhat standard Kripkean definition of model and of truth in a model for languages containing the modal logical constant 'A'.

[8] Zalta *identifies* the concepts of logical truth and of model-theoretic logical truth, so it is no surprise that he does not see his considerations as a criticism of any coextensionality thesis. But this view is obviously objectionable, for surely there is a wide conceptual gap between the notions of logical truth and model-theoretic logical truth. (This gap has been forcefully adverted to by Etchemendy (1990), even though Etchemendy's claims of non-coextensionality seem to me less fortunate.)

Now, a reasonable principle about logically true formulae is that any proposition they may come to express as a result of giving them an interpretation which respects the meanings of the logical constants ought to be necessary. However, if we interpret the letter ‘p’ by means of a sentence expressing a contingently true proposition, say “Kripke is a philosopher”, it appears that the content then expressed by (1) is contingent: in a possible world in which Kripke is not a philosopher it is still true that Kripke actually is a philosopher (because he is a philosopher in *our* world, the actual world), but in that world it is not true that he is a philosopher. This is reflected in the somewhat standard Kripkean model theory of the modal language we are considering, since the necessitation of (1),

(2) $\Box(Ap \supset p)$,

is not a model-theoretic logical truth; there is a model such that (1) is not true in some of the worlds of the model possible relative to α, and hence (2) is false at α in this model[9].

Thus we would supposedly have a model-theoretic logical truth in the sense of (T3), (1), which, when ‘p’ is interpreted in a suitable way, is not necessary. If we take it as a *datum* that a logical truth must be necessary (and this would seem eminently reasonable under most pretheoretical conceptions of the notion of logical truth), it follows that some model-theoretic logical truths in the sense of (T3) are not real logical truths.

One problem with this alleged counterexample is that it is a bit dubious that the actuality operator ‘A’ is a good candidate for logicality. But perhaps it is consistent with our vague intuitions about logical constancy to take ‘A’ to be a logical constant, or at least a potential logical constant. A more serious worry, however, is that although the model-theoretic definition of logical truth for modal languages containing the operator ‘A’ that Zalta uses is somewhat standard, it is not as entrenched as the corresponding definition for modal languages not containing ‘A’. In fact, there is room for choice concerning languages containing ‘A’ even within the *Kripkean* possible worlds semantics mentioned in (T3).

[9] David Kaplan (1977) gave an example very similar to (1) and claimed that it is a contingent logical truth. The example is ‘$ANp \supset p$’, where ‘N’ is a “now” operator with a semantics in tense logic analogous to that of ‘A’ in modal logic. Zalta’s example is less objectionable as a counterexample because ‘N’ is less clearly a logical constant than ‘A’.

Some theorists have adopted the following alternative. Call now a model a *quintuple* of the kind (W, R, α, β, V), where things are as before with W, R, α and V, and β is a member of W, possibly but not necessarily α. Keep all the recursive satisfaction clauses as before, but say that a sentence is true in a model (W, R, α, β, V) if it is true in β. Then clearly (1) is false in some models of the new kind, simply because it's false at some worlds in some models of the earlier kind.[10]

Zalta's examples would convincingly refute (T3) if they crippled all reasonable Kripkean model theories for languages containing 'A', or simply if they were directed at a more entrenched, or "the" entrenched Kripkean model theory. It would be good, of course, if there were fewer doubts about the choice of logical constants on which Zalta's counterexamples are based. But even if we grant that the choice is right, a suspicion remains because the examples are directed at the model theory of a non-extensional operator with a non-entrenched semantics. There is no standard recipe for constructing the Kripkean semantics of new modal operators we may come up with.[11]

In any case, as happened with (T1), one problem with (T3) and similar theses is that they talk about a very restricted set of logical constants. In fact, as long as one restricts oneself to a smallish finite set of logical constants (as in (T1), (T3) and similar theses), even the set of *extensional* logical constants among them may, for all we know, be always too restricted. And a problem peculiar to (T3), that in any case disqualifies it as a suitable Tarskian thesis, is that it talks about a notion of model (Kripke models) that Tarski simply does not talk about or even adumbrate in his classic paper of 1936.

To summarize, I think that the thesis that Tarski probably had in mind was something weaker than (T2) but not as specific as (T1) or (T3). Further, Tarski's thesis must have been one that seems reasonable to postulate when one restricts one's attention to classical propositional/quantificational models, for these are the models that Tarski clearly has in mind. Finally, Tarski's thesis must have made a broad claim about the *notion* of a logical constant—not about a set of logical constants characterized by specifying a mere list.

[10] For related critical remarks on Zalta, see Hanson (2006).

[11] Let me stress, however, that Zalta's example does succeed in refuting a particular coextensionality thesis, regardless of how reasonable or entrenched it may have been.

II

The following thesis, (T4), satisfies all these desiderata, and seems to me quite likely to capture the essence of what Tarski had, at least implicitly, in mind:

> (T4) A sentence of a formal language which possibly extends a classical propositional/quantificational language with new *extensional* logical constants which are propositional connectives, quantifiers and predicates is logically true in the pretheoretical sense iff it is true in all classical propositional/quantificational models which (re)interpret its constants (other than its extensional logical constants).

(T4) is just like (T2), but it restricts itself to quantificational languages with extensional logical constants, and tacitly presupposes a natural extension of the classical notion of truth in a model for such languages (see, e.g., the examples below). This was very probably Tarski's intent. There are well-known dismissive remarks of Tarski about the presumable impossibility of giving non-extensional constants "any precise meaning" (Tarski (1935), p. 161); these remarks are in his classic monograph on truth, published one year before his paper on logical consequence. Besides, it's clear that if he had had in mind a thesis about non-extensional logical constants, he would not have restricted himself to extensional models such as the classical quantificational models, in view of elementary considerations such as those that show that (T2) is false. Furthermore, (T4) seems to underlie to a good extent the practice of using classical models in the model theory of languages having as logical constants generalized quantifiers, the predicate of identity, and other extensional constants denoting notions invariant under permutations of the quantificational domain.

What does 'extensional' mean, exactly? It's hard to be precise, but the rough idea, which will be enough for my purposes, is this (the extension to the polyadic cases is obvious):

a) A (monadic) connective © is extensional if, whenever φ and φ' are formulae which are either both satisfied or both unsatisfied by the same valuation v of the variables at a world w, ©φ is satisfied by v at w iff ©φ' is satisfied by v at w.[12]

[12] Note that, as desired, 'A' is not extensional in this sense.

b) A (monadic) quantifier Q is extensional if, whenever φ and φ' are formulae which are satisfied at a world w by the same set of valuations differing at most at 'x' from a valuation v, $Qx\varphi$ is satisfied by v at w iff $Qx\varphi'$ is satisfied by v at w.

c) A (monadic) predicate P is extensional if, whenever t and t' are terms with the same denotation under a valuation v of the variables at a world w, $P(t)$ is satisfied by v at w iff $P(t')$ is satisfied by v at w.

Now that we know what (T4) means and having claimed that it probably deserves the title 'Tarski's thesis' more than any of the other theses we have considered,[13] we can ask: is (T4) true?

One thing that would seem clear is that (T4) is not trivially false. But there are a number of examples that have been given in the literature which, if they were convincing, would show that (T4) is false in a rather trivial way. Two of these examples are due again to Etchemendy (1990) and McGee (1992). Etchemendy has noted that, if one takes as logical constants the extensional monadic predicates 'P' and 'M', meaning respectively "is or was a president of the U.S. on or before 2005" and "is a male", then the quantificational sentence

(3) $(\forall x)(P(x) \supset M(x))$

is true in all classical quantificational models, since no matter what model we choose (3) will be true in the model. In this case, that means that no matter what (set-sized) quantifier domain of actually existing things we choose, every object in that domain will be either a non-president or a male (there haven't been any female presidents in the actual world). However, (3) is not a logical truth under most conceptions of logical truth, for it is not even necessary. Or, at least, (3) is not necessary if it is interpreted in such a way that it quantifies over any of a wide class of natural ranges for its quantifier; for example, if it ranges over "absolutely everything"[14], or over the set of humans, etc., then (3) is not intuitively

[13] As mentioned in an earlier note, in 1936 Tarski required the domain of some models to be the denotation of a non-logical predicate of the formal language under consideration. However, he later adopted the now common convention and presumably he stuck otherwise to his views in the 1936 paper.

[14] In the sense recently elucidated by Tim Williamson (2003).

necessary. And hence it is not a pretheoretical logical truth, under most conceptions of logical truth.

McGee has given another purported counterexample (but he has not categorically asserted that it is a counterexample). Take as a logical constant the extensional quantifier '$(\exists^{PC}x)$', meaning "there are at least a proper class of x's such that"; then the quantificational sentence

(4) $\sim(\exists^{PC}x)(x=x)$

is true in all classical quantificational models, since no matter what set-sized quantifier domain we choose, the sentence will be true in that domain. And yet (4) is not true for proper class-sized domains, much less necessary; or, at least, (4) is not necessary when the proposition it expresses quantifies over, e.g., "absolutely everything", or over the class of sets, etc. So (4) is not logically true, under most conceptions of logical truth.

A problem for anyone who wants to use these examples against (T4) is that the arguments needed for this must be premised on suspicious choices of expressions as logical constants. There is a patent intuition, I think, that neither 'P' nor 'M' are logical constants, and there is to say the least no clear intuition that '$(\exists^{PC}x)$' is a logical constant, so the persuasive force of Etchemendy's and McGee's examples is quite limited. This intuition also vindicates the initial impression that a refutation of (T4) cannot be trivial (or at the very least cannot be as trivial as the refutations that would be provided by those examples). Part of what makes finding a refutation of (T4) non-trivial is that one cannot choose or define just about any constants, pick any non-logical truth involving them and claim that that truth is a model-theoretic logical truth in the sense of (T4) because it is true in all models given that we decide to count the constants in question as logical. If one wants to refute (T4), one must rather find constants about whose logicality there is an independent and reasonably entrenched intuition, and show that some sentence that is not a logical truth is nevertheless a model-theoretic logical truth in the sense of (T4).

As I announced at the beginning, I do think that (T4) is false. But the reason why I think it's false is not trivial (at least not trivial in the just indicated sense in which the reasons provided by the preceding attempted counterexamples would be trivial if they worked). If I'm right that (T4) false, that will mean that a *prima facie* reasonable, non-trivial strengthening of (T1), that is in all probability the coextensionality thesis held by Tarski, is false, and hence that Tarski's theory of logical truth and logical consequence is defective to some extent.[15]

The way in which I will argue that (T4) is false will be by exhibiting certain quantificational sentences containing only constants that are intuitively logical and extensional, which are true in all classical quantificational models, but which, like Etchemendy's, McGee's and Zalta's sentences, are *not necessary* (in the sense that their non-logical vocabulary can be interpreted, and indeed it is typically interpreted, in such a way that the truths then expressed are not necessary). Since under most pretheoretical conceptions of logical truth, interpreted sentences that are not necessary are not logically true, it will follow that (T4) is false under such conceptions[16].

The examples I will present are based on an assumption about the modal behavior of the classical first-order quantifiers that is not part of their explicit classical model-theoretic semantics, but which is, I think, the most natural assumption one can make. There are basically two kinds of propositions that a formula dominated by a classical universal quantifier (a formula of the form '$(\forall x)\varphi$') might be taken to express, which are reflected in the two most standard semantics for quantificational modal logic. The first kind of proposition is a proposition the content of whose quantifier gets specified when one specifies a quantificational domain, given purely in extension, plus a property which further restricts the range of the quantifier at a particular world. Thus, for example, given this view, in order to specify the content of the quantifier in the proposition expressed by (a use of) the sentence (3), what one has to do is to specify a class of objects, given purely in

[15] As noted in section I, there are strong arguments for its adequacy when its range of application is restricted in the manner specified in (T1(1)) and even in (T1). In section III, however, we will see a possible qualification even of (T1(1)).

[16] On a minority of pretheoretical conceptions of logical truth, including perhaps Tarski's, a refutation of (T4) that appealed to intuitions about necessity would be dismissed on the grounds of some kind of skepticism about modality. However, I think it's most significant to evaluate (T4) using assumptions shared by the majority of views, and one such is the assumption that any bona fide logical truth must be necessary. Note that (T4) sounds initially plausible regardless of whether one is in the minority of skeptics about modality or in the majority of non-skeptics (just as (T1) has sounded plausible to many people regardless of the details of their pretheoretical conceptions of logical truth).

extension, plus a property. For example, one specifies the set {Franklin Roosevelt, Eleanor Roosevelt, George Bush, Laura Bush} plus the property of being Texan. Assuming that the other expressions have their intuitive meaning, (3) then expresses the proposition that, roughly, "each of the Texans in the set {Franklin Roosevelt, Eleanor Roosevelt, George Bush, Laura Bush} is, if a president, a male". This proposition is true in those worlds where, as in the present one, the Texan presidents in the set {Franklin Roosevelt, Eleanor Roosevelt, George Bush, Laura Bush} are all males. The way this idea is reflected in one of the standard semantics for quantified modal logic is as follows: given a previously specified domain D and a property Π, one says that $(\forall x)\varphi$ is satisfied by a valuation *v* at a world *w* iff every valuation *u* of the variables (with objects of the previously given domain D) which differs from *v* at most at 'x' and which assigns to 'x' an element of D that has the property Π in *w* satisfies φ at *w*. (Typically Π is taken to be the property of existence.) This clause has the effect that with a quantifier '$(\forall x)$' one quantifies in a world *w* only over the objects from D that have the property Π in *w*.

The second kind of proposition that a quantificational formula might be taken to express is a proposition the content of whose quantifier gets specified when one specifies simply a quantificational domain, given purely in extension, which constitutes the range of the quantifier at any particular world. Given this view, in order to specify the content of the quantifier in the proposition expressed by (a use of) the sentence (3), what one has to do is simply to specify a class of objects, given purely in extension. For example, one specifies the set {Franklin Roosevelt, Eleanor Roosevelt, George Bush, Laura Bush} as before, and that's enough. Assuming the other expressions have their intuitive meaning, (3) then expresses the proposition that, roughly, "each of the things in the set {Franklin Roosevelt, Eleanor Roosevelt, George Bush, Laura Bush} is, if a president, a male". This proposition is true in those worlds where, as in the present one, the presidents in the set {Franklin Roosevelt, Eleanor Roosevelt, George Bush, Laura Bush} are all males; and it will be false, for example, in worlds where Eleanor becomes president. The way this idea is reflected in the other standard semantics for quantified modal logic is as follows: $(\forall x)\varphi$ is said to be satisfied at a world *w* by a valuation *v* of the variables with objects of the previously given domain D iff every valuation *u* of the variables with objects of D which differs from *v* at most at 'x' satisfies φ at *w*. This clause has the effect that with a quantifier '$(\forall x)$' one quantifies in a world *w* over all the objects of the previously specified domain D.[17]

Note that under this way of understanding the content of the universal quantifier, when we consider whether, e.g., the quantificational sentence (3) is true at a world, we always ask ourselves whether each of the objects of a previously fixed domain D is either out of the extension of 'P' for that world or in the extension of 'M' for that world. Assuming that these extensions are the sets of presidents and males, respectively, what we always ask ourselves is whether each of *those same objects, the objects of that same domain* D, is either a non-president or a male in the world at issue[18].

Regardless of what of these two ways of understanding the modal behavior of the classical quantifier we choose, that does not conflict with its status as an intuitive logical

[17] One common way to illustrate the difference between the two just described semantics in a non-extensional language is by observing that the so-called Barcan formulae (e.g. $(\forall x)\Box P(x) \supset \Box(\forall x)P(x)$ or $\Diamond(\exists x)P(x) \supset (\exists x)\Diamond P(x)$) are true in all Kripke models under the second, "fixed domain" semantics but false in some models under the first, "variable domain" semantics. A quantificational Kripke model is a sextuple (D, W, R, α, V, Π), where things are as in the propositional models of section I with W, R and α; D is a non-empty set (intuitively of individuals); V an assignment of extensions drawn from D to each pair predicate letter-world; and Π an assignment of extensions drawn from D to each world (intuitively Π is the restricting property resorted to in the "variable domain" semantics, but idle in the other semantics). Take $\Diamond(\exists x)P(x) \supset (\exists x)\Diamond P(x)$. Under the fixed domain semantics, if $\Diamond(\exists x)P(x)$ is true in α, there is a world *w* R-accessible from α such that some object o of the fixed domain D is in the extension of P in *w*; but then o is an object of the fixed domain D such that there is a world *w* R-accessible from α, such that o is in the extension of P in *w*; so $\Diamond(\exists x)P(x) \supset (\exists x)\Diamond P(x)$ is true in α regardless of the model, and hence true in every model under the fixed domain semantics. On the other hand, to construct a counterexample to $\Diamond(\exists x)P(x) \supset (\exists x)\Diamond P(x)$ under the variable domain semantics, let $D=\{1,2\}$, $W=\{\alpha,w\}$, $R=W\times W$, $V(P,\alpha)=\varnothing$ and $V(P,w)=\{2\}$; also, let Π assign $\{1\}$ to α and $\{2\}$ to *w*. In this model, $\Diamond(\exists x)P(x)$ is true and $(\exists x)\Diamond P(x)$ false in α.

[18] Under this conception of the quantifiers, it is reasonable to postulate that a quantificational sentence has not been given a content unless a non-empty domain for the quantifiers has been stipulated. (This postulate provides a way out of the exception to Kreisel's intuitive soundness premise, provided by sentences like '$(\exists x)(P(x)\vee\sim P(x))$', mentioned in an earlier note.) In the same way, it is reasonable to postulate that a sentence containing indexicals has not been given a content unless context determines some object(s) as the denotations of the indexicals under the context. See below for more on indexicals and the quantifiers.

constant. Or, in the case of the first way, there is no conflict when the property Π is a logical property like existence. And in fact the two kinds of quantifiers are taken as logical constants in quantificational modal logic. Further, under both ways the universal quantifier is an extensional quantifier in the sense above (or rather, in the case of the first way, it is extensional provided the property Π is extensional).

I think it is most natural to suppose that the universal quantifier behaves modally in the second way just described, i.e. that with it one quantifies in every world over all the objects of the domain that serves to interpret the quantifier. (Since there is no property Π to worry about, the universal quantifier understood in this way is a logical constant without qualification.) The reason why I think this is most natural requires us to review quickly some well-known ideas from the philosophy of language.

In standard treatments, a sentence containing indexicals becomes truth-evaluable in a circumstance or possible world only under a context of use, a context of use which provides a content for the indexicals and indirectly for the whole sentence under the context. If I now utter the sentence 'I am a philosopher', this sentence, given that I am the speaker of the context, acquires under the context the content that Mario is a philosopher. A content of this kind can be evaluated for truth or falsehood in the present and other possible circumstances. It will be true in those circumstances in which Mario is a philosopher, and false in those circumstances in which Mario is not a philosopher. A similar view is widely held about indexicals which, unlike 'I' or 'actually', are not pure, such as demonstratives like 'that'. If I say

(5) That is a philosopher,

my utterance will have acquired a content under the context if the context has provided a content for the word 'that'. This content cannot have been provided merely by the linguistic rules for 'that' and my uttering the word, unlike what happens in the case of 'I'—this is what it means that 'I' and 'actually' are pure. Something else, perhaps my action of pointing toward some person, or my intending that the word 'that' indicate a certain person, must have happened. I will suppose that this "something else" happens, and won't go into what it might consist in. Suppose that my word 'that', as a result of some feature of the context of my utterance, came to indicate Kripke. Then the content of my utterance of (5) was that Kripke is a philosopher. This content will be true at those worlds in which Kripke is a philosopher, and false at those worlds in which Kripke is not a philosopher.

This way of dealing with indexicals and their content is encouraged by certain intuitions about the *rigidity* and the *direct referentiality* of indexicals. Recall what Kripke (1972) called rigidity. A designator is rigid if it designates the same object in all possible circumstances or worlds. (Well, actually Kripke's notion of rigidity is a bit weaker than this; the one I just defined is what Nathan Salmon (1982) christened with the specialized name of 'obstinate rigidity'.) One way in which one checks that a designator is rigid is by considering what would have to happen in order for the content of sentences containing that designator to be true in different possible circumstances. Kripke asks us to consider (6) and (7):

(6) Aristotle was fond of dogs.
(7) The last great philosopher of antiquity was fond of dogs.

When we consider whether the content of (6) is true in other possible circumstances, we always ask ourselves whether the same person, Aristotle, is fond of dogs in those circumstances. So Aristotle must be the object designated by 'Aristotle' when (6) is evaluated at a possible circumstance, and 'Aristotle' is rigid. On the other hand, when we consider whether the content of (7) is true in other possible circumstances, we ask ourselves whether the person who is the greatest philosopher of antiquity in those circumstances is fond of dogs in those circumstances, and this need not be always the same person, in particular it need not be Aristotle. So different objects can be designated by 'The greatest philosopher of antiquity' when (7) is evaluated at other possible circumstances, and so 'The greatest philosopher of antiquity' is not rigid.

Similar considerations strongly suggest that indexicals taken under a context are rigid. Consider (5) again. When we ask ourselves whether the content of (5) under the context (or, in alternative terminology often taken to be equivalent, what I *said* with my utterance of (5)) is true in other possible circumstances, we always ask ourselves whether the same person, Kripke, is a philosopher in those circumstances. So Kripke is the object designated by 'that' under the context when (5) is evaluated at a possible circumstance, and so 'that' is rigid.

Direct referentiality is the name David Kaplan gave to what he took to be a related property of names and indexicals. Perhaps the best way of understanding this property is by noting than certain definite descriptions are rigid. One example is ‘The ratio of the circumference of any circle to its diameter’. It designates the same object, the number π, in every circumstance. Consequently, when we consider whether the content of the sentence

> The ratio of the circumference of any circle to its diameter is worshipped by the ancient Babylonians

is true at a world, we always ask ourselves whether the same object, π, is worshipped by the ancient Babylonians in those circumstances. But although the description ‘The ratio of the circumference of any circle to its diameter’ is rigid, it is not directly referential. It has a descriptive content by means of which, or through the mediation of which, the same object turns out to get determined in all circumstances as the designation of the description. A name and an indexical (under a context), on the other hand, are (on Kaplan’s view) directly referential because their content is not such that it serves to “determine” the designated object; the content of a name or indexical (under a context) is directly the designated object, there is (in particular) no descriptive content conventionally associated with them which “determines” the designated object.

On reasonable assumptions about how to understand the classical first-order quantifiers, these would have properties analogous to rigidity and direct referentiality. These quantifiers are of course not designators in Kripke’s sense (they are not singular terms), but there is a relatively clear sense in which they make implicit “reference” to a set of objects over which the quantifier variables range, or to the objects themselves as a plurality. Given a domain of objects for the variables, sentence (3) says approximately that every thing in *that domain* is either a non-president or a male, or that every thing among *those things* is either a non-president or a male.

(3) does not contain any descriptive element which helps determine the domain of objects over which its variables range. Rather, that domain or those objects are given or fixed by the logician who uses a formal language containing (3) before (3) can be evaluated for truth and falsehood, in a way similar to the way in which something in the context provides a reference for ‘that’ in (5). Once the domain or its objects are fixed, (3) can be

evaluated for truth and falsehood. But (3) itself contains no descriptive element which determines that domain or its objects.[19] So it would appear that in some sense the quantifier '(∀x)' in (3), under a choice of a domain, can be taken to make implicit reference to that domain or its objects *in a direct way*. We may say that '(∀x)' *directly ranges over* its domain or its objects.

This proposal about the quantifier '(∀x)' ranging directly over its domain is in harmony with the standard extensional semantics for it. This semantics requires only that a set of objects be given for the range of the quantifiers to be determined. No descriptive property characterizing or restricting the objects of the set is demanded in interpretation. (This is especially appropriate in model theory, since the model theorist clearly does not want to leave out of his considerations models whose domain is not picked out by any descriptive property.) Further, the standard satisfaction clause for a formula of the form '(∀x)φ' says that (∀x)φ is satisfied by a model and a valuation *v* of the variables with objects of the domain of the model if every valuation *u* of the variables with objects of the domain which differs from *v* at most at 'x' satisfies φ. In this clause there is no mention of any descriptive property required to hold of the objects which can be assigned to 'x' (by valuations which differ from *v* at most at 'x').

Direct referentiality implies rigidity in the case of proper designators (but not vice versa, as we saw). Must the quantifier '(∀x)' be taken to be "rigid" if we decide to understand it as ranging directly over its domain? In other words: suppose '(∀x)' ranges directly over its domain, and we are considering the question whether the content of (3), provided by a choice of a domain D for its variables, is true in other possible circumstances; must we adopt the view that we always have to ask ourselves whether each of *those same objects, the objects of D*, is either a non-president or a male (in those circumstances)? Or briefly put: is the domain D the range of quantification of '(∀x)' in all

[19] Recall that in 1936 Tarski adopted the convention that the domains of the models for first-order languages ought to be denoted by a non-logical predicate of those languages, which in some cases could express a descriptive property. He further required quantifications in those languages to be relativized to that predicate. So the present argument would not quite work under Tarski's 1936 conventions; it would work only under his later conventions.

possible worlds? It appears so: assuming that (3), under a choice of a domain for its variables, makes implicit direct reference to that domain or its objects, and given that this domain or its objects are not determined by any descriptive element in the content of (3), it appears that the domain or its objects themselves must form part of the content of the quantifier. And then it appears that we must adopt the view that when we consider whether the content of (3) is true at a possible world, we always have to ask ourselves whether each of *those same objects, the objects of that same domain*, is either a non-president or a male (at that world). We thus have some powerful considerations in favor of the view that the first-order quantifiers behave modally so that the domain that serves to interpret the quantifiers is the domain over which one quantifies in every world.

An objection to this understanding of the classical first-order quantifiers as ranging directly and rigidly over their domain might run as follows: "It is true that the explicit textbook conventions about the first-order quantifiers don't say anything about whether they range directly over their domain or are "rigid". But a tacit assumption underlying what is explicitly said in textbooks is that the first-order quantifiers must be understood as being as close as possible to their correlates, or their closest correlates, in natural languages. Perhaps if we look at quantifiers in natural language we'll find that they are not "rigid", or do not range directly over their domain."

Now, I'm not sure that what this objection calls the "tacit assumption" of formalization is in fact such. As noted above, formalization focuses on the precisification of truth-conditional content, and does not care about whether features relevant to modal character are determined in the process. But let's concede the "tacit assumption" for the sake of argument. Are the quantifiers of natural language not "rigid", or do they not range directly over their domain? Any answer to this question is bound to be complex and controversial. In particular, the answer is bound to be complex and controversial in the case of quantifiers which work as determiners, such as 'every', 'any' or 'some' (in its use as a determiner).

But the process of formalization leaves it settled that first-order quantifiers do not work grammatically as determiners. Unlike, say, 'every', '($\forall$x)' does not grammatically work as a determiner of a noun or a predicate that restricts the reach of quantification, and requires only a formula in order for a formula dominated by it to get formed. First-order quantifiers seem to be closer to pronominal uses of quantifiers in natural language, as in sentences like (8), (9), and (10):

(8) All are seen in the evening.
(9) Some are seen in the evening.
(10) All are, if presidents, males.

I think that a good case can probably be made that pronominal quantifiers range directly over their domain and are thus "rigid". The traditional terminology which calls them 'pronominal' even leads us *a priori* to expect this, since it is common in the treatments of indexicals that I've been relying on to consider pronouns as indexicals, hence as "directly referential".

Imagine that my student Mary and I are looking at close-up photographs of Mars, Venus and Jupiter in an astronomy book. Then I utter (8). It may be useful to think of our intuitions about rigidity first. When considering whether the content of my utterance is true at a different possible circumstance, what will I consider? I think I will ask myself if each of these same objects, the objects in this domain containing Mars, Venus and Jupiter, are seen in the evening. What I said with my utterance will be true at a possible circumstance just in case all of Mars, Venus and Jupiter are seen in the evening in that circumstance. Something similar will happen in the case of (9). What I say if I utter (9) in the context just mentioned will be true in another possible circumstance just in case at least one among Mars, Venus and Jupiter is seen in the evening in that circumstance. What I earlier called an "implicit" reference to a domain seems to be taking place also here, and it seems to be taking place in a "rigid" way.

Do pronominal quantifiers range directly over their domain? Presumably. The only way that occurs to me in which "direct ranging" might not be taking place here in the presence of rigidity would be that pronominal quantifiers were in some sense disguised quantifier determiners whose complement noun phrase picks out rigidly a set of objects, but has been elided. For example, on this view (8) would be elliptical for (11), or (12), or something similar:

(11) All objects identical with either Mars, Venus or Jupiter are seen in the evening.
(12) All these are seen in the evening.

Arguably both ‘objects identical with either Mars, Venus or Jupiter’ and ‘these’ pick out the set B rigidly. Now, even if we conceded that the content of some such sentence was the real content of my utterance of (8), this would be no objection to the view that there was “direct ranging” involved, since on the views we are relying on ‘these’ is directly referential. But I don’t think one ought to concede even that. One basic problem with the suggestion is that there is no reason why my utterance of (8) should have the content of (11), or the content of (12), or the content of any one in particular of the many other options that we might think of. If there were in fact one noun phrase that I inadvertently elided when I uttered (8), presumably there ought to be some principle that determined what that noun phrase is and that allowed us to recover it; but there is no apparent principle that determines one such noun phrase as preferred.

Another, even more serious problem, lies in the fact that, even if there were such a principle determining a preferred noun phrase in some cases, it could hardly be seen as a principle determining indirectly the real content of my original utterance of (8) as the content of some other sentence. Presumably I can come to understand (8) before I understand ‘objects identical with either Mars, Venus or Jupiter’, or ‘these’, or other noun phrases that are candidates for ellipsis on the objector’s view. So, provided we grant the common idea that understanding is knowledge of content, one can know the content of my utterance of (8) before one knows the content of the sentences proposed by the objector; but on the objector’s view this would be impossible. Together with the underdetermination argument of the preceding paragraph, this argument about independent intelligibility seriously undermines the ellipsis objection.[20]

Having thus argued for what I think is the most reasonable way of understanding the modal behavior of the first-order quantifiers, let me go back to the promised counterexamples to (T4). The counterexample that is more usefully presented first involves the further reasonable premise that a monadic predicate ‘E’ meaning “exists” is a logical constant. A primitive predicate with this intended meaning is taken as a logical constant in treatments of quantified modal logic and of intensional logic generally (see, e.g., David

[20] Similar arguments against related syntactic ellipsis views of the quantifiers are provided by Stanley and Szabó (2000), though they don’t consider the case of pronominal quantifiers.

Kaplan's (1977), (1978) logic of demonstratives). Its intended meaning, a bit more explicitly, is given by the principle that it is to be satisfied by an object at a world if that object exists at that world. This is the common sense of existence according to which I exist in the current circumstance but I would not have existed if a certain spermatozoid and a certain egg had not met. It can be given a simple satisfaction clause in the definition of satisfaction in a classical quantificational model, very much like the clause for identity: a model M plus valuation v satisfies E(x) if v(x) exists.[21] Clearly 'E' is an extensional predicate in the semi-formal sense above. (In the sense that when t and t' are terms with a shared denotation under a valuation v at a world w, E(t) is satisfied by v at w iff E(t') is satisfied by v at w.)

Is the predicate 'E' really a logical constant? As I just said, it has been taken to be such in intensional logic. But furthermore 'E' is certainly topic-neutral. Moreover, it (or its correlate in natural language) is widely applied across different areas of discourse, and thus satisfies the main criterion for logical constancy mentioned by defenders of "pragmatic" conceptions of this notion (see e.g. Hanson (1997), Warmbrōd (1999), Gómez-Torrente (2002)). Arguably, 'exists' is used in mathematics, for example; or at the very least, it is unclear that every appearance of the word as applied to mathematical objects is to be understood as an appearance of something like the first-order or a higher-order existential quantifier. Further, 'E' is a logical constant in the technical senses defined by Tarski, McGee and Solomon Feferman. It is surely invariant under permutations of a model (Tarski (1966), Tarski and Givant (1987)) and even under bijections of models (with a domain of existing things), for its extension in any model (with a domain of existing things) is the full domain of the model; and it follows from its meaning that it is invariant under bijections of models (with a domain of existing things) (McGee (1996)); it is also invariant under homomorphisms of models (with a domain of existing things) in the sense recently defined by Feferman (1999). It is not a logical constant in the technical sense of Timothy McCarthy (1981)[22], since it is not invariant under bijections of models with a domain of existents plus

[21] Note that 'E(x)' is not equivalent to 'AE(x)' under the somewhat standard model theory for 'A' used by Zalta, since an object may exist in a world without existing in the actual world.

[22] Nor in the related sense of Gila Sher (1991), (1996). Sher has emphasized to me that her domains contain non-existents. In this sense her proposal seems to me quite non-Tarskian.

possibly non-existing things; but this would seem a defect of McCarthy's proposal rather than a virtue.[23]

(It is worth mentioning that (T4) is easily refuted if one understands 'logical constant' not in a pretheoretical way as I do, but by means of any of the "permutationist" notions mentioned in the preceding paragraph (including McCarthy's), as I have argued in Gómez-Torrente (2002). Observe that a predicate 'H' meaning "is a married bachelor" has an empty extension in all possible worlds. Given this, it is invariant under bijections of all models without qualification; yet it is not a logical constant, provided there is any distinction at all between logical and non-logical constants. But suppose for the sake of argument that 'H' is a logical constant, as a "permutationist" characterization would have it. It follows that a sentence like '$(\forall x){\sim}H(x)$' is a model-theoretic logical truth in the sense of (T4). But clearly this sentence is not a logical truth. Tarski (in (1936)) presumably understood the notion of logical constancy pretheoretically, assuming at most that invariance under permutations (of actual domains) is a necessary condition on logical constancy. In what follows we will see that, under most conceptions of logical truth, (T4) is false even thus understood.)

Now consider the following quantificational formula[24]:

(13) $(\forall x)E(x)$.

Given only the reasonable premise that 'E' is a logical constant, this formula is a model-theoretic logical truth in the sense of (T4). Since there are no non-logical constants to worry about, a classical quantificational model for (13) is just a (set-sized) quantifier domain composed of existing things. And no matter what classical quantificational model for (13) we choose, (13) will be true in the model: no matter what actual (set-sized)

[23] And a defect of the related proposal of Sher (1991), (1996) (see the preceding note).

[24] I am taking the quantifier '$(\forall x)$' in (13) to be the classical first-order quantifier, and assuming that it behaves modally as in the "fixed domain" semantics described above—in short, that it is "rigid". For the purpose of giving a counterexample to (T4), however, one doesn't need to assume that the quantifier in (13) *is* the classical first-order quantifier. One may simply (and uncontroversially) assume that it is the "rigid", extensional and logical quantifier described above, leaving aside the question whether it is to be identified with the usual first-order universal quantifier.

quantifier domain composed of existing things we choose, every object in that domain will be an existing thing (in our world). This presupposes the usual understanding of quantificational model theory, according to which no non-existing objects form part of the domains of models.[25]

On the other hand, given only our earlier conclusion that the first-order universal quantifier appearing in a formula like (13) is "rigid", (13) can easily be interpreted so as to express a contingent content, and thus is not a logical truth under most conceptions of logical truth. Suppose that we specify the domain of the quantifier in (13) to be again the set {Franklin Roosevelt, Eleanor Roosevelt, George Bush, Laura Bush}. Under this interpretation, (13) is true in the actual world, but only contingently: in a possible world in which Franklin's parents had not met, (13) would not be true, because at least one of the people it quantifies over (they are the same four people in all possible worlds) would not exist in that world.[26] The same would happen with any domain containing a contingent existent.[27]

So given that 'E' is an extensional logical constant, and on the reasonable assumption we defended earlier, that the first-order universal quantifier is "rigid", (13) is true in all classical quantificational models (that reinterpret only the constants other than extensional logical constants), is actually true in many domains for its quantifier, and, furthermore, for many of those domains it expresses a contingent truth. But under most pretheoretical

[25] Etchemendy's example (3) (and other examples of his), as well as most discussions of these issues, are based on this presupposition. If the presupposition is not made, the example does not clearly work. There may well be a non-existing person who is president of the United States and a female, even in our world—perhaps some fictional characters are non-existents with this property.

[26] For related reasons, Kaplan (1977) argued that 'I exist' is a (model-theoretic) logical truth (in his "logic of demonstratives") which is not necessary. For our purposes this is irrelevant, since 'I', unlike the quantifiers, is presumably a non-logical constant.

[27] Under some natural interpretations, (13) is even false; for example, if we take the domain of the quantifier to consist of "absolutely everything", (13) is false provided only that "something" (in the absolutely unrestricted sense) is not an actual existent. The same happens with any more restricted domain containing a non-existing object. However, I do not want to make my argument against (T4) rely on these interpretations, for the question whether it is possible to quantify over non-existing objects is substantially controversial. (Though I myself see no problem with it in the case of some mere *possibilia*.)

conceptions of logical truth, a sentence that is a logical truth cannot be interpreted so as to express a non-necessary truth. It follows that under most pretheoretical conceptions of logical truth, thesis (T4), what I think was Tarski's thesis, is false.

The falsity of (T4) under most conceptions of logical truth does not depend on the fact that 'E' is a logical constant, for arguably other sentences analogous to (13), but not containing 'E', are model-theoretic logical truths in the sense of (T4) but can be interpreted so as to express contingent truths. A general lesson of (13) is that, for any monadic predicate *F* that is intuitively an extensional logical constant, applies to all existing things, but fails to apply to a thing in worlds where it doesn't exist, the sentence '$(\forall x)F(x)$' will be a model-theoretic logical truth in the sense of (T4), but it will not be a pretheoretical logical truth, under most conceptions of logical truth. It seems clear that other predicates with these properties exist. For example, a dyadic predicate '$T(x, y)$' meaning "is a part of" is intuitively extensional, and has often been called a logical constant. Arguably the monadic predicate '$T(x, x)$' applies to all existing things, but presumably fails to apply to a thing in worlds where it doesn't exist[28]. Another example with the same features as '$T(x, y)$' may be a predicate '$S(x, y)$' meaning "is simultaneous with" (in the sense of "it occupies the same stretch of time as"[29]); also arguably, '$S(x, x)$' possesses the same relevant features as '$T(x, x)$'. Given these assumptions, the sentences

$$(\forall x)T(x, x)$$

and

$$(\forall x)S(x, x)$$

[28] The latter is again a frequent assumption about atomic predicates in modal logic. (Nevertheless, I doubt that it's a reasonable assumption for *all* such predicates. Elsewhere (Gómez-Torrente (2006)) I have argued that it may not be an intuitively compelling assumption about a number of atomic predicates whose application to an object at a world does not seem to depend on the object's existence. In my view, this is what happens with the predicate '$x = x$' of self-identity, with the predicate 'x is a thing', and with predicates expressing a natural kind to which the object belongs essentially (such as 'x is a cat' or 'x is an electrical discharge').)

[29] Note that this predicate intuitively applies even to abstract, hence presumably eternal existents.

will be further examples of sentences which are model-theoretic logical truths in the sense of (T4) but are not real logical truths under most conceptions of logical truth.

III

The examples in section II are based on a natural assumption about the modal behavior of the classical first-order quantifiers, suggested by features of their classical extensional semantics. As I announced at the beginning, the examples in this final section will be based on more speculative assumptions about the epistemology of the quantifiers, which are nevertheless also suggested by features of their classical extensional semantics. Just as the standard semantics of the classical quantifiers is not explicit about their modal behavior, it is also not explicit about some aspects of their semantics relevant to their epistemology. In this section I will explain how, given some plausible, but potentially more controversial ways of understanding the epistemology of the quantifiers, one obtains the result that some model-theoretic logical truths in the sense of (T4), and in fact even of (T1) and (T1(1)), are not *a priori*, and hence presumably not logically true.

We argued in section II that the classical first-order quantifiers are most naturally understood as ranging directly over their domain and thus as "rigid" (though only the rigidity of the quantifier appearing in (13) was needed as an assumption for the argument that (13) is not logically true). The thesis that the quantifiers range directly over their domain suggests a familiar idea for finding model-theoretic logical truths in the sense of (T4) (and even of (T1) and (T1(1))) which are not *a priori* knowable. Before developing it in an explicit way, some clarifications.

What can it mean to say that a *sentence*, like (3), is *a priori* knowable? The question whether a sentence is *a priori* knowable does not seem to be a felicitous question unless it is made against the background of some suitable assumptions. A formula or even a sentence is not the sort of thing one can know or fail to know. Let's then ask the question about the *content* we gave to (3). Is it *a priori* knowable? Even this second question appears not to be proper. The reason is that it is unclear what has to happen for me to bear some psychological attitude toward the content of an expression containing directly referential or "directly ranging" expressions, like the content of (3). In fact, it is not clear that my bearing some psychological attitude toward something can be the same as my bearing some simple relation to what we have been calling a content.

The problem emerges in several ways. Think first of this famous example:

(14) Hesperus is Phosphorus.

Kripke noted that if ‘Hesperus’ and ‘Phosphorus’ are rigid, then the content of (14) is necessary given that it is true. He also noted that, intuitively, what we might call the (typical) attitudinal content of (14) is knowable only *a posteriori*. But if proper names are directly referential, (15) has the same content as (14), given the way we have been speaking of “content”:

(15) Hesperus is Hesperus.

And yet the (typical) attitudinal content of (15) would seem knowable *a priori*. If we accept these intuitions at face value, it follows that the attitudinal contents of (14) and (15) must be different, even if their contents are the same. (Sometimes these attitudinal contents are called ‘cognitive contents’.) But the case of (14) and (15) leaves open the possibility that the sentence is a relevant part of its attitudinal content; on this view, (14) and (15) might have different attitudinal contents simply because they are different sentences.

However, both Kripke (1979) and Kaplan (1977) have noted the possibility that different occurrences of the same sentence have the same content but different attitudinal contents. This can happen if different occurrences of the same directly referential word in the same sentence have the same content and yet differ in attitudinal content. For example, different occurrences of a demonstrative may share content but be accompanied with different demonstrations which carry distinct cognitive imports. Think of a weird but otherwise unobjectionable utterance of (16) where the utterance of the first ‘that’ is made on an evening, pointing to the Evening Star, and the rest of the utterance is made some later (much later) morning, pointing to the Morning Star:

(16) That is that.

This utterance of (16) has the same content as (14) and (15) (under the assumption that names and demonstratives are directly referential), and given our description of the conditions of the utterance, it appears that in some sense its attitudinal content is knowable only *a posteriori* (like that of (14)). On the other hand, if (16) is uttered quickly in the evening, with the speaker accompanying the two utterances of (16) with an action of continuously pointing to the Evening Star, it seems that the attitudinal content at stake can in some sense be known *a priori*. The two utterances of (16) would share sentence and content but would differ in attitudinal content.[30]

Here is another way in which the problem emerges, now illustrated directly for the case of the first-order quantifiers when these are taken to range directly over a domain. What can it mean to say that one knows (whether *a priori* or not) a quantificational sentence dominated by a quantifier '$(\forall x)$'? If we answer that it is to know its content, the content that is directly signified, there is still a problem. For in what sense can I be determinately said to know of *the objects* in the set, that all have a certain property? Suppose the domain of the variables is the set B={Mars, Venus, Jupiter}, and that the predicate 'SIE(x)' means "is seen in the evening", so that every utterance of

(17) $(\forall x)$SIE(x)

has roughly the content that all things in the *set* B are seen in the evening. One perfectly acceptable way of introducing B as the domain for the intended interpretation of (17) is for me to utter (17) while pointing to Mars, Venus and Jupiter in the evening sky. If I do this it appears that in some sense the attitudinal content of my utterance is not very informative. People can be assumed to have known it already. But if I utter (17) while pointing to close-

[30] If proper names are directly referential, I suspect that (14) and (15) can have more than one attitudinal content, and, even in the case of (15), attitudinal contents that are *a posteriori*. I suspect that in this respect the case of proper names is more similar to that of demonstratives than often realized. It would seem perfectly possible that different tokens of a name like 'Hesperus' are associated by a speaker with different attitudinal contents, even if they are tokens with the same semantic content. (Compare the Paderewski cases brought out by Kripke (1979).) These tokens might appear in the same sentence. If (15) can have attitudinal contents that are *a posteriori*, even if its tokens of 'Hesperus' have the same semantic content (Venus), then we would have an argument that (15) is not a logical truth, under the assumption made below in the text that a logical truth cannot have *a posteriori* attitudinal contents. Yet the straightforward first-order formalization of (15), '$a=a$', is a model-theoretic logical truth. (Compare a related but different remark in Salmon (1986), p. 176, n. 5.)

up photographs of Mars, Venus and Jupiter in the astronomy book (another perfectly acceptable way of introducing B as the domain of quantification), it seems that the attitudinal content at stake is informative, and people who trust me can learn a great deal from my utterance.

So if we accept all these intuitions at face value, it appears that sentences like (16) and (17) taken together with their contents are not (or not simply) what one is related to when one can be said to hold the relevant psychological attitudes, for example the attitude of *a priori* knowledge toward the attitudinal content of (an utterance of) (16) or the attitude of having learned toward the attitudinal content of (an utterance of) (17).

If a sentence and even a sentence taken together with a content for it are not determinately *a priori* or *a posteriori*, when should we call a sentence *a priori* (or *a posteriori*)? One possibility is to call a contentful sentence *a priori* when *some* of the attitudinal contents it may have are *a priori*. But this may lead to a trivialization of the notion of apriority for true sentences containing quantifiers. Think of me stipulating to my student John that the "rigid", directly ranged over domain of the quantifiers is "the set of planets seen in the evening in the coordinates such and such", and suppose with me, for the sake of argument, that this set is the set B above. I synthesize the application of this stipulation to formulae starting with a universal quantifier by means of

(18) $(\forall x)\varphi \equiv$ For every x which is a planet seen in the evening in the coordinates such and such, φ.

The two sides of this biconditional of course differ in content, but since (18) is a stipulation, John can perhaps be said to know its attitudinal content *a priori*. And then John can also perhaps be said to know *a priori* that

(17) $(\forall x)SIE(x)$,

even if the content of (17) is contingent. (This is basically the argument used by Kripke in his famous example of the sentence 'One meter is the length of the standard bar in Paris'.) (17) would be a contentful sentence with at least one *a priori* attitudinal content. It seems thus that if logical truths are to have an epistemological property which distinguishes them from (17), this property ought not to be the property of having some *a priori* attitudinal content.

So let's opt for the other alternative that immediately suggests itself: let's agree to call a contentful sentence *a priori* when *all* the attitudinal contents it can have are *a priori*. (This principle is adopted by Salmon (1986).) Then we ought to accept this reasonable principle about interpreted logically true formulae: any attitudinal content they may come to have as a result of giving them whatever is needed to give them an attitudinal content ought to be an *a priori* knowable attitudinal content. If we find an interpreted model-theoretic logically true formula which is susceptible of having an attitudinal content that is knowable only *a posteriori*, it will follow from this principle that some model-theoretic logical truths are not *a priori*, and thus are not real logical truths under typical conceptions of logical truth.

Is (13), when one stipulates the range of its quantifier to be, e.g., {Franklin Roosevelt, Eleanor Roosevelt, George Bush, Laura Bush}, only susceptible of having an *a posteriori* knowable attitudinal content? One is tempted to say yes, since in a moderately intuitive sense one ought not to be able to know *a priori* that Franklin Roosevelt existed, and this would seem to be part of what one would have to know *a priori* in order to know *a priori* any attitudinal content expressed by (13) with its content. But, at any rate, it might seem clear that (13) with its content is susceptible of having at least *one* attitudinal content which is knowable only *a posteriori*. Think of someone uttering (13) while pointing to the portraits of Franklin Roosevelt, Eleanor Roosevelt, George Bush and Laura Bush. Presumably whatever attitudinal content is involved in this situation is one I can only be said to be able to know *a posteriori*.

But here is a worry about the alleged aposteriority of an attitudinal content of (13). Perhaps in order for me to be able to quantify over some objects these objects must exist in my world. Perhaps this principle about quantifiability is knowable *a priori* by me. If this is so, then it is reasonable to think that I will be able to know *a priori* that any attitudinal content expressed by (13) with its content is true. Perhaps it is then also reasonable to think that I will be able to know *a priori* any attitudinal content expressed by (13) with its content. (13) with its content would then be one more alleged example of the contingent *a priori*.

I propose to forget about (13) in connection with apriority. Given our assumption that the first-order quantifiers are directly referential, in an utterance of a quantificational sentence containing two or more quantifiers it appears that we ought to find the same phenomenon as in (16). For definiteness, think of (19):

(19) $(\forall x)SIE(x) \supset (\forall x)SIE(x)$,

where the quantifiers are interpreted as ranging again over the set B. (19) is a model-theoretic logical truth, even if 'SIE' is not a logical constant. Is (19) with its content knowable *a priori* or only *a posteriori*?

Again the answer would appear to be that that's not an appropriate question. Mere sentences containing expressions that signify their contents directly, or even sentences of this kind provided with a content, are not what one is related to when one can be said to know something. Before the question is asked, more needs to be said about the circumstances in which one comes into contact with (19) and its content.

Suppose that I utter (19) assertively in front of John. At the same time that I utter each token of the quantifier in (19), imagine that I point to Mars, Venus and Jupiter in the evening sky. I think in this situation it can properly be said that John has not learnt anything that he could not have found out by himself *a priori*.

Now suppose that Mary also knows that the domain B is the range of the variables, and that I utter (19) assertively in front of her. At the same time that I utter my first token of the quantifier in (19), imagine that I point to Mars, Venus and Jupiter in the evening sky, as before. But at the same time that I utter my second token of the quantifier, imagine that I point to the close-up photographs of Mars, Venus and Jupiter in the astronomy book. I think that in this situation it can properly be said that Mary has learnt *a posteriori* something that she could not have found out by herself *a priori*.

It is certainly possible to introduce the domain of quantification for a first-order quantifier by any of the means we have mentioned, and so it is very reasonable to think that different tokens of a first-order quantifier having the same content differ in attitudinal content. What might raise more doubts is the claim that, once the domain has been introduced in some way, different tokens of a quantifier within the same sentence might have different attitudinal contents. Would it be more reasonable to think that, once the domain has been introduced in some way, some tacit principle prevents us from assigning different attitudinal contents to different tokens of a quantifier within the same sentence, so that something has gone wrong in my description of the attitudinal content of my utterance of (19) to Mary? I think not. The directly ranged over domain of quantification in the intended interpretation for a language often receives different descriptions from the person who introduces it, for example in a textbook or during a class. These descriptions often convey different descriptive information that gives rise to different attitudinal contents, and nothing seems to prevent these descriptions from being given in the time interval between the inscriptions of different tokens of a quantifier in the same sentence. The following situation seems perfectly natural: I am teaching a class and I stipulate that the domain of quantification of the sentence I am about to write on the blackboard is the set B above, which I introduce by pointing to three photographs of Mars, Venus and Jupiter as tiny spots seen in the evening sky, but I don't call the planets by their names; then I write the antecedent of (19), but before writing the consequent I note that the objects in B are precisely the objects called 'Mars', 'Venus' and 'Jupiter' in some astronomy book containing close-up pictures of the planets, a book that my students know well; I then finish writing (19). It is most reasonable to think that the attitudinal content that the first token of '$(\forall x)$' has for my students is different from the attitudinal content of the second token, and that (19) has for them an *a posteriori* attitudinal content.

The standard stipulations about the semantics of classical quantificational languages include the stipulation that all occurrences of a quantifier in a sentence are to range over the same initially fixed domain. This stipulation is respected in the examples. John, Mary and my other students knew, just by knowing this and the other stipulations, that (19) had to be true, perhaps they can even be said to know this fact about the sentence *a priori*. But by means of my utterance of (19) Mary learnt something knowable only *a posteriori* while John did not learn anything he could not have learnt *a priori*. If we are not too fussy about using "word salads" mixing natural and formal language, such as are frequent in logic and mathematics texts, it even appears that we may use (19) embedded in attitudinal clauses to report what Mary learnt *a posteriori* and John could have learnt *a priori*. Imagine utterances of the following sentences, accompanied by actions of pointing similar to the ones I described earlier:

Mary learnt *a posteriori* that $(\forall x)SIE(x) \supset (\forall x)SIE(x)$.
John could have learnt *a priori* that $(\forall x)SIE(x) \supset (\forall x)SIE(x)$.

As noted in section II, someone might embrace the idea that the first-order quantifiers must be understood as being as similar as possible to their closest correlates in natural languages, and object to the counterexample claiming that different utterances of a quantifier of natural language in the same sentence are not susceptible of having different attitudinal contents. However, just the opposite seems true. It appears that pronominal quantifiers have an epistemology essentially identical to that of indexicals. If I utter (8) or (9),

(8) All are seen in the evening,
(9) Some are seen in the evening,

neither the sentence nor the sentence taken together with its content seem appropriate objects for psychological attitudes. Something more is needed for an attitudinal content to arise. If I utter (8) to Mary while pointing to Mars, Venus and Jupiter in the evening sky, intuitively she cannot be said to have learnt something she (we may assume) did not know. If I utter (8) to her while pointing to Mars, Venus and Jupiter in the astronomy book, intuitively she can be said to have learnt something she (we may assume) did not know. Also, if I utter

(20) If all are seen in the evening then all are seen in the evening

(the sentence that would seem to provide the most accurate translation of (19) into English), accompanying the utterance of the first 'all' with an action of pointing to Mars, Venus and Jupiter in the evening sky and the utterance of the second 'all' with an action of pointing to Mars, Venus and Jupiter in the astronomy book, then Mary can be said to have learnt something she could not have learnt *a priori*.

Suppose one argued (as in the related ellipsis protest of section II) that the two pronominal quantifiers in the relevant utterance of (20) are disguised quantifier determiners with two elided complement noun phrases that in this case have the same content but different attitudinal contents. It may be good to note that this by itself would be no objection to the claim that there are model-theoretic logical truths that are not *a priori*. Presumably those noun phrases would be syntactically identical, as in an utterance of

(21) If all these are seen in the evening, then all these are seen in the evening

(which we may imagine made accompanying the utterance of the first 'all these' with an action of pointing to Mars, Venus and Jupiter in the evening sky and the utterance of the second 'all these' with an action of pointing to them in the astronomy book). But even if the relevant apparent utterance of (20) (and indirectly of (19)) is not really an utterance of (20) (or (19)) but is elliptical for (21), we still have reason to think that it is an utterance of a model-theoretic logical truth. After all, the sentence that it is supposed to be elliptical for, (21), may be translated back to formal language as a sentence of the form '$(\forall x)(P(x) \supset Q(x)) \supset (\forall x)(P(x) \supset Q(x))$', with antecedent and consequent having the same truth-conditional content. So this version of the ellipsis objection is no obstacle to the claim that there are model-theoretic logical truths that are not *a priori*.

A different version of the objection would be to claim that the two pronominal quantifiers in the relevant apparent utterance of (20) (though perhaps not in other utterances of (20)) are disguised quantifier determiners with two elided *syntactically different* complement noun phrases that have the same content but different attitudinal contents. For example, my described apparent utterance of (20) would on this view be elliptical for an utterance of (22), (23), or something similar:

(22) If all objects identical with either Mars, Venus or Jupiter are seen in the evening, then all objects identical with either Mars, Phosphorus or Jupiter are seen in the evening.

(23) If all these are seen in the evening, then all those are seen in the evening.

In this way, it might be argued, an appearance is created that (20) (and indirectly (19)) can have *a posteriori* attitudinal contents, though in reality this appearance is created by the fact that, in the circumstances of the relevant apparent utterance of (20), it is elliptical for (22), (23), or some other similar not logically true sentence, which can in fact have *a posteriori* attitudinal contents. However, even assuming for the sake of argument that the relevant apparent utterance of (20) is in fact elliptical for something, it's hard to see why it should be elliptical for an utterance of a sentence with two syntactically different complement noun phrases instead of a sentence with two occurrences of the same noun phrase. For example, an utterance of (21) is fully intelligible if I accompany the utterance of the first ‘all these’ with an action of pointing to Mars, Venus and Jupiter in the evening sky and the utterance of the second ‘all these’ with an action of pointing to them in the astronomy book. Why should my utterance of (20) be elliptical for (23) and not for (21) (if it’s elliptical for anything at all)?

In fact, however, I think it’s reasonably clear that my utterance of (20) is not elliptical for any sentence of the sort, and that this is shown by underdetermination and independent intelligibility arguments analogous to those of section II. There is no apparent principle determining that my apparent utterance of (20) should be elliptical for (21), or for (22), or for (23), or for any one in particular of the many other options that we might think of. And even if some such sentence was determined as preferred for some syntactic or psychological reason, it could hardly be seen as a required intermediate step in the determination of the real content of my utterance. Intuitively, this utterance (and presumably any felicitous utterance of (20)) is intelligible, and hence has a knowable content, independently of a previous knowledge of the meaning and hence the content of ‘objects identical with either Mars, Phosphorus or Jupiter’, ‘these’, ‘those’, and other noun phrases that are candidates for ellipsis on the objector’s view. Intuitively, we can easily imagine vicissitudes in which Mary does in fact not know the meaning of those noun phrases but my utterance of (20) is still perfectly intelligible to her, and makes her learn something she could not have learnt *a priori*. On the objector’s view such independent intelligibility, and hence Mary’s ability to learn something *a posteriori* with just an independent understanding of my utterance of (20), are impossible, against the intuitive view.

I conclude that (19) with its content is neither *a priori* nor *a posteriori* by itself. But (19) with its content, and any interpreted model-theoretic logical truth containing more than one occurrence of a quantifier, will have some attitudinal contents which are *a priori* and some which are *a posteriori*. If we accept the principle mentioned above, that all the attitudinal contents of an *a priori* sentence with its content must be *a priori*, then all the attitudinal contents of a logical truth with its content must be *a priori*. It follows that some model-theoretic logical truths in the sense of (T4), and even of (T1) and (T1(1)) (in fact many of them) are not real logical truths under typical conceptions of logical truth, which require logical truths to be *a priori*.

A reasonable hypothesis is, however, that all attitudinal contents of model-theoretic logical truths of unexpanded first-order quantificational languages in which the attitudinal contents of the different tokens of the quantifiers are all the same will be *a priori*. Thus a reasonable further constraint, that presumably turns (T1) and (T1(1)) into true theses, is the condition that all the attitudinal contents of the quantifiers must be the same.[31] Then the essence of the preceding considerations is that, if the reasonable (but not uncontroversial) assumptions of this section are accepted, some such constraint on the widely accepted (T1) and (T1(1)) is needed.

References

Bays, T. (2001), "On Tarski on Models", *Journal of Symbolic Logic*, vol. 66, pp. 1701-1726.

Church, A. (1936), "An Unsolvable Problem of Elementary Number Theory", *American Journal of Mathematics*, vol. 58, pp. 345-363.

Cook, S. (1971), "The Complexity of Theorem-Proving Procedures", in M.A. Harrison, R.B. Banerji and J.D. Ullman (eds.), *Proceedings of the third annual ACM symposium on Theory of computing*, ACM Press, New York, pp. 151-158.

Cook, S. (1991), "Computational Complexity of Higher Type Functions", in I. Satake (ed.), *Proceedings of the International Congress of Mathematicians in Kyoto*, Springer, New York, pp. 55-69.

Etchemendy, J. (1988), "Tarski on Truth and Logical Consequence", *Journal of Symbolic Logic*, vol. 53, pp. 51-79.

Etchemendy, J. (1990), *The Concept of Logical Consequence*, Harvard U.P., Cambridge (Mass.).

Feferman, S. (1999), "Logic, Logics and Logicism", *Notre Dame Journal of Formal Logic*, vol. 40, pp. 31-54.

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

Gómez-Torrente, M. (1998/9), "Logical Truth and Tarskian Logical Truth", *Synthese*, vol. 117, pp. 375-408.

Gómez-Torrente, M. (2000), "A Note on Formality and Logical Consequence", *Journal of Philosophical Logic*, vol. 29, pp. 529-539.

Gómez-Torrente, M. (2002), "The Problem of Logical Constants", *Bulletin of Symbolic Logic*, vol. 8, pp. 1-37.

Gómez-Torrente, M. (2006), "Rigidity and Essentiality", *Mind*, vol. 115, pp. 227-259.

[31] Perhaps a similar further constraint about the attitudinal contents of individual constants is also required, if individual constants are directly referential. Recall the comment about '*a*=*a*' in the preceding footnote.

Hanson, W. (1997), “The Concept of Logical Consequence”, *Philosophical Review*, vol. 106, pp. 365-409.

Hanson, W. (2006), “Actuality, Necessity, and Logical Truth”, *Philosophical Studies*, vol. 130, pp. 437-459.

Kaplan, D. (1977), “Demonstratives”, in J. Almog, J. Perry and H. Wettstein (eds.), *Themes from Kaplan*, Oxford U.P., New York, 1989, pp. 481-563.

Kaplan, D. (1978), “On the Logic of Demonstratives”, *Journal of Philosophical Logic*, vol. 8, pp. 81-98.

Kreisel, G. (1967), “Informal Rigour and Completeness Proofs”, in I. Lakatos (ed.), *Problems in the Philosophy of Mathematics*, North-Holland, Amsterdam, 1967, pp. 138-171.

Kripke, S. (1972), “Naming and Necessity”, in D. Davidson and G. Harman (eds.), *Semantics of Natural Language*, Reidel, Dordrecht, 1972, pp. 253-355, 763-769.

Kripke, S. (1979), “A Puzzle about Belief”, in A. Margalit (ed.), *Meaning and Use*, Reidel, Dordrecht, pp. 239-283.

McCarthy, T. (1981), “The Idea of a Logical Constant”, *Journal of Philosophy*, vol. 78, pp. 499-523.

McGee, V. (1992), “Two Problems with Tarski’s Theory of Consequence”, *Proceedings of the Aristotelian Society,* n.s., vol. 92, pp. 273-292.

McGee, V. (1996), “Logical Operations”, *Journal of Philosophical Logic*, vol. 25, pp. 567-580.

Mancosu, P. (2006), “Tarski on Models and Logical Consequence”, in J. Ferreirós and J. J. Gray (eds.), *The Architecture of Modern Mathematics*, Oxford University Press, Oxford, pp. 209-237.

Quine, W. V. (1954), “Quantification and the Empty Domain”, *Journal of Symbolic Logic*, vol. 19, pp. 177-179.

Ray, G. (1996), “Logical Consequence: a Defense of Tarski”, *Journal of Philosophical Logic*, vol. 25, pp. 617-677.

Salmon, N. (1982), *Reference and Essence*, Blackwell, Oxford.

Salmon, N. (1986), *Frege’s Puzzle*, M.I.T. Press, Cambridge (Mass.).

Sher, G. (1991), *The Bounds of Logic*, M.I.T. Press, Cambridge (Mass.).

Sher, G. (1996), “Did Tarski Commit ‘Tarski’s Fallacy’?”, *Journal of Symbolic Logic*, vol. 61, pp. 653-686.

Soames, S. (1999), *Understanding Truth*, Oxford U.P., New York.

Stanley, J. and Z.G. Szabó (2000), “On Quantifier Domain Restriction”, *Mind and Language*, vol. 15, pp. 219-261.

Tarski, A. (1935), “The Concept of Truth in Formalized Languages”, in Tarski (1983), pp. 152-278.

Tarski, A. (1936), “On the Concept of Logical Consequence”, in Tarski (1983), pp. 409-420.

Tarski, A. (1966), “What Are Logical Notions?”, *History and Philosophy of Logic*, vol. 7, 1986, pp. 143-154.

Tarski, A. (1983), *Logic, Semantics, Metamathematics,* second edition, Hackett, Indianapolis.

Tarski, A. and S. Givant (1987), *A Formalization of Set Theory without Variables*, American Mathematical Society, Providence, R.I.

Turing, A. (1936/7), “On Computable Numbers, with an Application to the Entscheidungsproblem”, *Proceedings of the London Mathematical Society*, ser. 2, vol. 42, pp. 230-265.

Warmbrōd, K. (1999), “Logical Constants”, *Mind*, vol. 108, pp. 503-538.

Williamson, T. (2003), "Everything", in D. Zimmerman and J. Hawthorne (eds.), *Philosophical Perspectives 17: Language and Philosophical Linguistics*, Blackwell, Oxford, pp. 415-465

Zalta, E. (1988), "Logical and Analytic Truths that Are not Necessary", *Journal of Philosophy*, vol. 85, pp. 57-74.