Title:

Tarski and Carnap on Logical Truth - or: What Is Genuine Logic?

Author:

Gerhard Schurz

IPS-PREPRINTS

Annual 1998 No. 1

Edited by Gerhard Schurz und Alexander Hieke

Vorveröffentlichungsreihe am Institut für Philosophie der Universität Salzburg

Prepublication Series at the Department of Philosophy, University of Salzburg

!Attention: In this Preprint all formulas are set into Latex Code ($......$)*!*

Tarski and Carnap on Logical Truth - or: What Is Genuine Logic?

Gerhard Schurz, University of Salzburg

1. Introduction

I came to this topic along my logical investigations of the Is-Ought problem in multi-modal logics (Schurz 1997). There are infinitely many mathematically possible modal logics. Are they all philosophically serious candidates? Which modal logic the "right" one -- does such a question make sense? A similar question can be raised for the infinite variety of propositional logics weaker than classical logics. The *Vienna Circle* concept of logic was that logic holds merely by *form*, *independent* from the facts of the world. Have we lost this concept completely? Is it a matter of arbitrary choice, of mere subjective-practical appropriateness, which logic one chooses? Is Quine right that there is no distinction between analytic and synthetic truth, even not, if we take "analytic truth" in the narrow sense of "logical truth"? These are the questions which have motivated this paper.

Tarski as well as Carnap have tried to define a sharp borderline between logical versus extralogical truths, and logical versus extralogical concepts. Both attempts are complementary in several respects. To some extent, this paper can also be seen as describing a game between Tarski and Carnap, which we call the Tarski-Carnap-game, in short the T-C-game. Section 2 constitutes Round 1 of the T-C-game and goes to Tarski, section 4 is Round 2 and goes to Carnap, section 7 is Round 3 and goes again to Tarski, finally section 8 is Round 4 which goes to Carnap. So the Tarski-Carnap-game ends 2:2, draw - which is as it has to be, of course. The other sections of the paper reflect on the contemporary discussion on 'genuine logic' and present some own suggestions.

2. Tarski's Definition of Logical Truth (Round 1 of the T-C-Game)

In (1936), Tarski presented for the first time his new *semantic* definition of logical consequence and logical truth. He starts to motivate his definition by critizicing traditional syntactic definitions. He gives two reasons why syntactic definitions are not satisfactory. *First*, the traditional *calculus-based* syntactic definitions (defined by a set of axioms and rules, and a recursive notion of proof) are *too weak* to capture the ordinary notion of logical consequence. Tarski gives the example of ω-incomplete theories of Peano arithmetics. Here it may be the case that $P(n)$ is derivable (within the theory) for every natural number n, without that $\forall x P(x)$ is derivable, although the latter sentences is intuitively following (Tarski 1936,

p. 410f). Tarski continues that this is just one aspect of the incompleteness of first order arithmetics, which shows that calculus-based recursive definitions are generally too weak to capture the intuitive notion of logical truth and consequence - and Tarski refers to Gödel's famous (1931) paper.

Second, Tarski discusses another possibility of a syntactic definition of logical truth which rest on the idea of closure under all substitution instances - and idea which goes back to *Bolzano* (Tarski 1936, pp. 415-417). This idea presupposes a *substitutional* interpretation quantifiers (as opposed to their referential interpretation) - i.e., $\forall x Fx$ is defined to be true iff all substitution instances Fa (for any individual term a of the language) are true. Here, Tarski mentions the German edition of Carnap (1937), who also used this substitutional interpretation (as we shall see, even in his 1947). The substitution interpretation presupposes that every object of the domain under consideration is denoted by an individual term of the language - an assumption which often is not satisfied, e.g., if the domain contains the real numbers. And whenever this assumption is not satisfied, then a substitutional approach will lead to intuitively wrong results. Hence, as Tarski concludes, also this second sort of syntactical definition does not capture the intuitive notion of logical truth and consequence. Yet Bolzano's and Carnap's syntactic attempts are close to Tarski's semantic approach, and Tarski suggests that his approach can be viewed as a semantic reconstruction of certain ideas of Bolzano and Carnap (Tarski 1936, p. 417).

Tarski sets up two *adequacy requirements* for his explication: the concept of logical truth and consequence should, first, be a *formal* concept, i.e., should be independent from the reference of certain nonlogical terms like "red", "bachelor" etc. (1936, p. 414f), and second, it should grasp the meaning of *necessary* (or *analytic*) truth, or implication (1936, p. 414, 419).

Tarski's approach has been intensively discussed and criticised by Etchemendy (1990). I follow Etchemendy in that I will focus on *logical truth*, while Tarski focuses on *logical consequence*. For our purpose, the difference is completely minor - all considerations about logical truth apply similarly to logical consequence, and vice versa (see below).

Tarski's approach is intended as one which holds for *every* kind of formal language. Of course, one should not quantify over "every" language. It is sufficient to assume that our approach should be applicable to the following formal languages: that of propositional logic, 1st order logic, higher order logic, and also modal logic - which will be discussed only in section 10.

Tarski as well Etchemendy (1990, pp. 32f) sharply distinguish between several *kinds of symbols* of a language. First, there is the distinction between the set of *variables* **Var** and the set of *primitive terms* **Term**. This terminology comes from Etchemendy; in what follows I will mean with "term" always "primitive term".

Primitive terms are often also called "constants" (e.g. individual constants), but as we shall see, this terminology is misleading. In the first order language, **Var** includes only the individual variables, while **Term** includes *all other symbols* (except brackets, which we don't count as extra symbols) - i.e, function symbols (including individual 'constants' as 0-ary function symbols), predicate symbols (including propositional 'variables' as 0-ary predicates - attention: they are *not* variables in our sense), and the set of logical symbols $\{\neg, \vee, \forall\}$ plus the defined symbols $\wedge, \to, \exists$, etc.

Formulas are called *sentences*, if they do not contain free variables, otherwise they are called *sentential functions*. Sentences are *made true or false* by *models*, or *interpretations*. In contrast, sentential functions are neither true nor false. Rather, they are *satisfiable* by *variable assignments*, or *sequences*, as Tarski calls them: these are functions which assign to each free variable of the sentence, and moreover to every variable of the language, an extension of the appropriate type based on the domain.

Tarski's definition of a *possible interpretation* presupposes a second and more critical distinction – a division of all terms into two disjoint subsets: the set **V** of *non-logical terms* – Etchemendy calls them *variable terms,* and a set **F** of *logical terms* – Etchemendy calls them *fixed terms.* Is important not to confuse the notion of a variable with that of a variable term. The usual division in standard first order language is to identify **F** with $\{\neg, \vee, \forall\, =}$ plus bound variables plus defined logical symbols, and **V** with the functions and relation symbols (except $=$). While the interpretation of the terms in **F** *remains fixed* (under all "possible" interpretations), that of the terms in **V** may vary arbitrarily, provided the *type-requirements* are met. For instance, the interpretation of a propositional operator or connective is a constant truth function; the interpretation of a quantifier is a constant set-theoretic function (see sections 7,9), but the interpretation of an individual 'constant' is an arbitrary object, and that of a n-ary predicate is an arbitrary set of n-tuples of objects.

Etchemendy presents three equivalent versions of Tarski's definition of logical truth. The common feature of all three definitions is that they are *relative* to a presupposed distinction between fixed and variable terms (if the set **F** is given, the set **V** is determined as **Term**–**F**). It is *this* distinction with help of which the intuitive notion of a *possible interpretation* is precisely defined (see D2 below).

The first definition (D1) is that of Tarski (1936, pp. 416f; see also Etchemendy 1990, pp. 45-47). It is assumed that the language contains for each type of term a corresponding type of variable. Hence, an extension of the language of 1st order logic by predicate variables (but not by 2nd order quantifiers) is needed to express (D1) for 1st order **L**-truth. If S is a sentence, then $S[x / \phi: \phi \in {\bf V}]$ denotes the result of replacing all variable terms in S by variables of the same type. Hence $S[x / \phi: \phi \in {\bf V}]$ will contain predicate variables where S contains predicate 'constants', and so on. "**L**-true" abbreviates "logically true".

(D1) S is **L**-true [w.r.t. **F**] iff $S[x / \phi: \phi \in {\bf V}]$ is satisfied by all sequences (variable-assignments).

In the second definition (Etchemendy 1990, pp. 53-56) one avoids the replacement of variable terms by variables and directly assigns the (varying) extensions to variable terms. Hence no language extension is needed to express this definition. Let us call a *model* or an *interpretation* a function which assigns to each fixed term in **F** its fixed extension and to each variable term in **V** an arbitrary extension which is constructed from the domain and belongs to the appropriate type. Then:

(D2) S is **L**-true [w.r.t. **F**] iff S is true in all models (interpretations).

(D2) corresponds 'roughly' to the modern standard model-theoretic definition of **L**-truth - but not *exactly*, as we will see soon.

In the third version, the arbitrary variation of the interpretations of variable terms is linguistically expressed by introducing higher order quantifiers - e.g., for expressing **L**-truth in 1st order logic, we need here a 2nd order language. Let $\forall x: S[x / \phi: \phi \in {\bf V}]$ denote the universal quantification of $S[x / \phi: \phi \in {\bf V}]$ with respect to all of its variables. Then:

(D3) S is **L**-true [w.r.t. **F**] iff $\forall x: S[x / \phi: \phi \in {\bf V}]$ is true.

All three definitions are equivalent, or at least, they *should* be equivalent in the 'right' conception of logic. This follows from Tarski's remarks, and it is explicitly pointed out by Etchemendy (1990, pp. 95-100). However, from the modern conception of models as pairs $\langle Dm,I \rangle$ - where Dm is the *domain* and I the *interpretation function* - one may point out that there is a difference between (D1,2) and (D3). For in (D1,2) we quantify in the metalanguage also over *all domains*, not only over all interpretation functions, while in (D3) we quantify only over all interpretations functions, so the truth of $\forall x: S[x / \phi: \phi \in {\bf V}]$ in (D3) will depend on Dm. But Tarski never speaks of a varying domain, he just speaks of the *class of all infinite sequences* taking arbitrary objects or sets as their components. This is equivalent with saying that Tarski *assumes a given constant domain* containing all set-theoretical objects one is allowed to consider in the meta-language. Given this view, then the three definitions are *indeed* equivalent.

3. John Etchemendy's Critique: What are Logical Symbols?

Etchemendy first points out that Tarski's definition, being relative to a given selection **F** of fixed terms, yields intuitively wrong results for unusual selections (1990, pp. 60 - 64). If, for instance, the sentence "snow is white" is included in **F**, then "snow is white" would come out as **L-**true, while if the operator $\neg$ would be treated as variable term (interpretable by any kind of unary truth function), then even "snow is white or not white" would not be **L**-true. Prima facie, this observation of Etchemendy is rather trivial, for Tarski's definition presupposes, of course, the *right* selection of **F** – that one in which **F** contains all and only *logical* terms. *But what is the criterion for the right selection?* This is Etchemendy's *main challenge*. He thinks that this distinction is arbitrary (1990, ch. 9). But even if one does not think so – one must be admit that as long as this distinction is based merely on *convention* instead on a clear philosophical criterion, the complaint of arbitrariness can not be refuted.

It should be emphasized that Tarski himself has clearly seen this problem. He considers it as the "perhaps most important open problem" (1936, p. 418) of his account to give a clear criterion for the standard distinction between logical and nonlogical terms. Tarski also asserts the following: if we would identify **F** with **Term**, i.e., consider all terms as fixed, then logical consequence would coincide with material implication, and logical truth with (material) truth (1936, p. 419). Indeed, this is immediately implied by (D1), since if **F = Term**, then **V** is empty, whence $S[x / \phi: \phi \in {\bf V}]$ coincides with S, so (D1) says in this case that S is logically true iff S is contingently true. This claim of Tarski is important because it expresses Tarski's belief that for every *constant* sentenc*e* - i.e., every sentences which contains no variable terms - logical truth coincides with contingent truth. This will be important for one of our adequacy criteria for 'genuine' logic.

Etchemendy's next example is the sentence "all bachelors are males". Intuitively, this sentence is 'necessarily' true. But the sentence is not **L**-true according to Tarski's definition, given the 'usual' selection of **F**, since predicates like "bachelor" and "male" are treated as variable terms. Standing alone, this argument does not constitute an objection to Tarski's definition. For as we have said, Tarski required the notion of **L**-truth to be a *formal* notion, independent from the particular meaning of predicates like "bachelor" or "male". It seems that Tarski's account is correct and all what is missing is a deeper justification of the usual distinction between logical and nonlogical terms. But here comes Etchemendy's *deepest* argument. The characteristic feature of a sentence like "All bachelors are male" is that it is true merely by virtue of the *meaning* of its terms – true independently of what the real facts are. According to philosophical standard terminology, such sentences are called *analytically* true (1990, p. 103). Now, Etchemendy argues as follows. Independently of how the standard selection of **F** is justified – if Tarski's account is adequate, then it should be an explication of the notion of *analytic truth w.r.t.* ***F***, i.e. of truth merely because of the meaning of the fixed terms. If we take the standard selection of **F**, Tarski's account should coincide with **L**-truth in the standard sense, while if we include predicates like "bachelor" and "male" in **F**, it should give us a wider notion of *analytic truth w.r.t. certain predicates*. Can we conceive Tarski's definition in this way? As Etchemendy demonstrates (pp. 105f, pp. 126-129), the answer is clearly "no". Compare the two sentences:

(1) All bachelors are males.
(2) All (U.S.) presidents are males.

Intuitively, (1) is analytically true, while (2) is contingently true. *But Tarski's definition does not reflect this difference.* If we include the predicates "(U.S.) president" and "male" in **F**, then also the second sentences comes out as **L**-true, because in every model which only varies the interpretation of terms other than $\{\neg, \forall, \to, president, male\}$, the sentence (2) will be true. But the truth of (2) is clearly world-dependent.

We learn from Etchemendy's example that Tarski's definition cannot be an adequate account for analytic truth in general. For nonstandard selections of fixed terms it produces wrong results. One might still hope that at least for the standard selection of **F** Tarski's definition yields the right results. Etchemendy thinks that these hopes are illusionary. He argues that the failure of Tarski's definition lies much deeper – namely in the fact that Tarski's definition reduces **L**-truth to contingent truth. Such an attempt can never succeed, for it will always make the notion of **L**-truth world-dependent and thus is ultimately inadequate (Etchemendy 1990, ch. 9). Etchemendy argues that this is not only true if **F** is selected in an unusual way, but even for the usual selection of **F** in 1st order logic, and he tries to substantiate this claim with arguments about quantifiers to be discussed in section 5 and 9. However, here I do not agree.

4. Carnap's Conception of Logical Symbols and Logical Truth (Round 2 of the T-C-Game)

Etchemendy is right that Tarski's definition is inadequate as a general definition of analytic truth. But this does not imply that it also must fail as a definition of **L**-truth. To infer this conclusion, a second premise is necessary: Etchemendy's claim that the distinction between logical and nonlogical terms is *arbitrary*. My goal is to develop a definition of logical terms which shall not only solve the problem of distinguishing in a non-arbitrary manner between fixed and variable terms, but will simultaneously *explain why* Tarski's definition of **L**-truth – conceived as "analytic truth w.r.t. **F**" – is adequate if and *only if* **F** contains only and all logical terms in the proper sense. My starting point are certain *ideas of Carnap*, mainly exposed in his (1947) and to some extent already in his (1937).

First two remarks on Carnap in general. Carnap's attempt to give a general characterization of logical symbols and logical truth may seem to contradict his *principle of tolerance* with respect to language systems, but it really does *not*. He sets up this principle (1937, p. 51) in the same book where he firsts attempts to define logicalness in an precise way (1937, p. 177-179). His idea seems to be that, of course, there may be different language systems with different logical rules, but in all these systems, logicalness and logical symbols are characterized in the same way (because this characterization is *independent* from particular rules, see below). Thus, Carnap's tolerance principle is restricted - it must *not* be understood in a Feyerabend sense of "anything goes". Second, compared to Tarski, Carnap is philosophically much broader, but as I see it, logically less precise and sometimes incoherent. In particular in his (1947), Carnap is guided by certain general ideas, which he captures in his "conventions", and later on he tries to state precise explications, in his "definitions", but while his "conventions" seem to me basically right and fruitful, his "definitions" seem to me misleading and sometimes even *wrong*, on reasons which I will explain in section 6. So I focus here on Carnap's ideas, i.e. on his general "conventions", and not on his "definitions". Carnap gives two of them (additions in square brackets are by myself):

"*Convention 17-1:* A designator is **L**-determinate [is a logical symbol] in [a semantical system] S iff its extension can be determined on the basis of the semantical rules of S alone, without any reference to [extra-linguistic] facts" (Carnap 1947, p. 70).
"*Convention 2-1:* A sentence is **L**-true in a semantical system S iff its truth can be established on the basis of the semantical rules of S alone, without any reference to [extra-linguistic] facts" (Carnap 1947, p. 10).

I rephrase Carnap's idea in this way, and call it the *criterion C:*

(C) A symbol is logical if its extension follows logically from the semantical postulates and rules of the metalanguage alone.

One might argue that there is a circularity: we already presuppose the concept of logical consequence in the metalanguage. In general, this is a deep point: every semantic characterization presupposes some amount of logic already in the metalanguage. But for our purpose of distinguishing logical from nonlogical symbols in the object language, this meta-logic reference is harmless and *non-circular*. *For*, even if we include intuitively extralogical postulates like "bachelors are male" etc. in the metalogic, this enriched metalogic will still *not* imply a determined extension of "bachelor" and "male". On the other hand, the ordinary semantical rules in the ordinary metalogic *will* imply a determined extension for each ordinary logical symbol, as we shall see. So the metalogic-reference will *not* lead to a relativity of the underlying distinction.

The semantical rules or postulates for terms are exactly what characterizes the *intension* of these terms, as opposed to their extension. Thus, there is the following alternative way of rephrasing Carnap's idea:

(C*) A symbol is logical iff its intension logically determines its extension.

Carnap's criterion fits obviously the standard logical terms of propositional logic: their intensions, i.e. their definitions, are the statements of their *truth tables*, which in turn are identical with, or at least completely determine their extensions, namely the truth-functions which they denote. As we shall see later, the same holds for quantifiers, if we reconstruct them in the proper sense. On the other hand, (C) is obviously false for all standard nonlogical terms, like bachelor, or male, and so on: no set of semantical postulates or rules will logically determine their extension, i.e., will tell us for all space and time points in the universe which living things are male and which are not male, etc.

Carnaps *criterion* of logical symbols *explains why* Tarski's definition of **L**-truth coincides with "analytical truth w.r.t. **F**" if and only if **F** contains all logical terms of the language. Tarski's definition deals only with the *extensions* of terms: in logically true sentence, only the extensions of logical terms are fixed while the extensions of all nonlogical terms may vary arbitrarily. So we can say that in Tarski's account exactly those sentences are logically true where the truth depends only on the *extension of their fixed terms*. But given the *right* selection of **F**, namely those terms which obey Carnaps criterion, then the extension of **F**-terms will be logically determined by their intension. Hence, if the **F**-terms satisfy the Carnap criterion, then the truth of logically true sentences in the Tarski sense will indeed be completely determined by the *intension* of their fixed terms, and hence will indeed coincide with *analytical* truth w.r.t. **F**. Thus, combined with the Carnap criterion, Tarski's account seems to be right.

Carnaps criterion gives also a clear division of the set of all analytic truths into *logical truths* and *extralogical analytic* truths. Extralogical analytic truths follow from *meaning postulates* for nonlogical terms, i.e. terms having a variable extension which is not determined by their intensions. Such meaning postulates correspond to what Etchemendy (1990, p. 71f) calls *cross term restrictions* on the range of possible interpretations of variable terms. For instance, the meaning postulate "bachelors are males" imposes the cross term restriction I(bachelor) $\subseteq$ I(male) on the range of admissible interpretations. Of course, cross term restrictions violate the Tarskian condition that logical truth should be a formal concept - *independent* from particular meaning of variable terms and hence be closed under arbitrary substitution for variable terms. Thus, we obtain the following first criterion for a 'genuine' logic - we call it the *independency criterion I:*

(I) In a 'genuine' logic the interpretations of variable terms can be varied independently of each other.

The distinction between logical and extralogical analytic truths based on the criterion C has also a bearing on the general debate about the *analytic-synthetic distinction*. An extralogical meaning postulate like "x is an A iff x is a B and x is a C", though *true by convention*, may still be more or less *empirically adequate*. One condition for its empirical adequacy is that the complex property "Bx and Cx" is indeed realized in the real world, even stronger, that it plays some significant role. Many of the Quinean objections against the analytic-synthetic dichotomy concern this sort of empirical adequacy. We may give up a definition because it is no longer empirically adequate. The same point *cannot be made* for genuine logical meaning postulates about logical terms which satisfy the criterion C. Here, the extension is completely determined by their intension - it does not depend on the world, we need not look onto the world to see whether the extension of a logical term is realized in the world. Therefore, a definition like "${p \leftrightarrow q}$ iff $p \to q$ and $q \to p$" cannot be called more or less empirically adequate in this sense. In general, the distinction between logical truths and extralogical truths (whether analytic or synthetic) seems to be much more stable and profound than the distinction between extralogical analytic and synthetic truths. If this is true, then a major part of the contemporary controversy about the analytic-synthetic distinction was on the wrong track.

Tarski's and Carnap's accounts are *complementary*. Quite ironically, Carnap *had* the means of distinguishing between logical and extralogical analytic truths - at least the preliminary means contained in his "conventions" - but he did *not* state this distinction. On the other hand, Tarski wanted a clear distinction between logical and extralogical analytic truth, but he did not have the means for it.

5. The Puzzle of Quantifiers

Does Tarski's account backed up by Carnaps criterion now fit the standard view of logical terms for 1st order logic? To our surprise, the answer is still *no*. The obstacle are the quantifiers. In the usual view, quantifiers as well as *bound variables* are counted as *logical* terms. Of course, also the identity sign is counted as a logical terms. Hence, sentences without individual and predicate terms (except identity) are *constant* and thus should be **L**-determinate (either **L**-true or **L**-false). Not so in standard 1st order logic. Consider *cardinality assertions*, like "There are exactly n objects", defined in the usual way in 1st order logic. This cardinality assertion is a constant sentence which is true iff the domain has cardinality n. Hence it is not **L**-determined in (standard) 1st order logic, although it *should* be **L**-determined according to Tarski's account.

The Tarskian account of **L**-truth gives right results *only* if every variable parameter of the semantic models for language can be identified as the variable interpretation of some variable term. This is violated for the standard view of 1st order logic - and as we shall see, it is much stronger violated for modal logics. Thus, we arrive at a second criterion for a logic in the 'genuine' sense - we call this the *transparency criterion T.* Logics satisfying T are transparent in the sense that their syntax fully reflects their semantics.

(T) In a 'genuine' logic, every variable semantical parameter must be the variable interpretation of a variable term of the language.

In 1st order logic we have the domain Dm as a varying semantic parameter which is not assigned to any variable term. In modal logic, we have a set of worlds W and an accessibility relation R as varying semantic parameters which is not assigned to any variable term.

6. A Modification of Carnaps Account to Logical Truth

Tarski (1936) was not explicit about possible variations of the domain, and the question of **L**-truth of 1st order cardinality assertions. How deals Carnap with this problem? The answer is maybe surprising: for Carnap , all cardinality assertions are indeed logically determined. How come?

This is the point where we have to make some comments about what in our opinion goes wrong in Carnap's attempts to replace his general ideas (conventions) by exact definitions. For instance, Carnap's definition of **L**-truth is this (1947, p. 10, 2-2.): *A sentence is* ***L****-true iff it holds in every state description*. This seems to be not very closely related to Carnap's convention 2-1 for **L**-truth stated in section 4. What goes on here?

Carnaps state descriptions are what in modern logic is called *diagrams*; they contain, for every atomic sentence of the language, either the sentence or its negation. The relation of "a sentence holds in a state description" is defined syntactically in recursive fashion (1947, p. 9). Carnap's state description approach presupposes i) that there is one *constant* domain Dm, ii) that each object in Dm has exactly one *standard name*, (iii) that quantifiers are interpreted substitutionally (in the sense of section 2), and finally iv) that standard names are viewed as **L**-determinate designators (i.e., as logical symbols) *by stipulation* (1947, p. 9f, pp. 73ff). A predicate is defined as **L**-determinate iff all its full instances with standard names are **L**-determinate (1947, p. 83).

Carnap calls his languages *coordinate languages*. He thinks that such coordinate languages occur not only in mathematics, but also in physics and more generally, in objectual disciplines where individual terms refer to real-world objects (1947, p 74). As a result, all identity statements with standard names and all cardinality assertions come out as **L**-true in Carnap' account. Standard names like "the North pole" and predicates like "x is identical with the North pole" are logical predicates in Carnap's account. I think that all this is extremely *counterintuitive*. The extension of real object names like "the North pole" will always depend on *our contact to the real world*. The intension of real object names, i.e. their purely language-internal description, can never *logically* determine their reference object - this determination will *always* be dependent on *contingent* facts of the world - think on the well-known examples of the morning and the evening star, etc. Moreover, Carnap's view of logically determined real object names violates Tarski's formality condition: "North pole" is not a purely formal concept, not an "armchair" concept, so to speak, which can be defined independent from any contact to the real world.

One might object: well, this argument only shows that the semantical rules do not only consist of language-internal, but also of language-external rules, like ostensive "definitions", etc. But note that this was not Carnap's view - according to him, all semantical rules are language-internal.[1] And this language-internality of semantical rules is the heart the separation of logical from extralogical terms. So I propose to rescue Carnap's criterion by distinguishing between *language-external* versus *language-internal* semantical rules. External rules are mainly ostensive rules like

---

1 Why, then, did Carnap hold his strange view? Maybe on two reasons. One the one hand, he defined the **L**-determinateness of predicates with help of the **L**-truth of their full sentences, and for this purposes he *needed* to assume that there exist at least some **L**-determinate individual terms. On the other hand, the early Carnap (1928, pp. 16ff, 204ff) held indeed the strange view that the reference of linguistic symbols to the world could be determined by purely language internal means, by purely structural descriptions ("structural maps"), without any ostension or other sort of world-contact. In this respect, the early Carnap was more a "universal logicist" than an "empiricist". However that may be, the fundamental mistake here is that even *if* a purely structural description would be sufficiently detailed to pick out in the whole universe exactly that unique spot on exactly that unique planet earth which we call the Northpole, then this determination of reference would *still* not follow logically, but merely contingently, from such a description, because the real universe could be different, in a way were the reference would *not* be uniquely determined by the "structural map" (the intension).

"this is the north pole", or "this is an example of water", they presuppose a particular perceptional *contact* with the world. In contrast, the internal rules, e.g. rules for the truth tables of propositional connectives, do *not* presuppose any particular (perceptional) contact to the world. With this distinction I suggest to sharpen Carnap's criterion as follows:

(C^+) A symbol is logical iff its extension follows logically from the purely language-internal semantical postulates and rules of the metalanguage.

This modified criterion seems to yield the intuitively right results. According to (C^+), no individual constant denoting a real individual can ever be a logical symbol. In contrast, several predicates are purely logical, e.g., the identify predicate $=$, or the universal predicate $Fx \vee \neg Fx$, etc. Quantifiers, if reconstructed properly, will also turn out as logical symbols. It might be objected that the modified Carnap criterion (C^+) is *intensionally* formulated. Is a purely extensional rephrasement of (C^+) possible? It seems yes, if we presuppose a purely *algebraic* formulation of the semantics of the language, where extensions are assigned to *all* terms, fixed or variable. But how could this be possible, if the domain of objects comes from the world? The trick is: simply by defining the extensions of logical terms as unique *functions*, going from arbitrary domains into set-theoretical structures over that domain. How this can be done is indicated in section 9. Given such an algebraic semantics, then the extensional version of the modified Carnap criterion would simply state that a symbol is logical iff its extensions remains the same in all interpretations which are admitted by the language-internal rules.

The only difficult point is our distinction between language-external and language-internal rules. This distinction may be not always sharp, at least not on "logical reasons". In any case, this distinction seems to involve an *unavoidable intensional aspect*. Is there possibility of replacing the criterion (C^+) by a criterion which avoids notions like "language-internal rules" and can be reformulated in a purely extensional way - "in purely Tarskian spirit", so to speak? This question brings us to the most important *alternative suggestion* for logicalness of terms, which has been developed by the master himself some decades later.

7. The Invariance Condition (Round 3 of the T-C-Game)

The basic idea of the invariance criterion (IV) is this:

(IV) A symbol is logical iff its extension is invariant under arbitrary permutations of the domain.

A permutation is a 1:1.mapping of the domain onto itself. For instance, an n-ary predicate F is logical iff for all objects $a_1, \ldots, a_n$ and permutations $p: Dm \to Dm$, $\langle a_1, \ldots, a_n \rangle$ is in $I(F)$ (the extension of F) iff $\langle p(a_1), \ldots, p(a_n) \rangle$ is in $I(F)$. If F is unary, this immediately entails that F can only be logical if F's extension is either empty (the "contradictory" predicate) or identical with the domain (the "tautological" predicate). More details soon; first something about the history of this criterion. The invariance idea for logicalness was first coming up in the paper of Tarski and Lindenbaum (1934/35), where the authors prove that every relation expressible by purely logical symbols in a Russellian type-theoretic logic is invariant under such permutations. Fritz Mautner (1946) then proposed this criterion as a criterion for logical symbols. Later, Mostowski (1957) suggested this criterion as one to distinguish logical from nonlogical quantifiers. In 1966, Tarski gave a lecture - published after his death as Tarski (1986) - where is describes this criterion very clearly (as always) and adopts it as his favoured criterion for logical symbols. Recently, Gila Sher wrote a book (1991) where she defends this invariance criterion against the challenge of Etchemendy (1990) and describes it at length.

Tarski (1986) motivates the invariance criterion as a generalization of Felix Klein's "Erlanger Program" for geometrical concepts. Here, purely geometrical concepts had been defined by geometrical invariance conditions. First take an *Euclidean geometry* - structurally, a set of three-dimensional points, a so-called vector space, which obeys the Euclidean axioms. Certain things are *arbitrary* in an Euclidean coordinate system: for instance, the choice of the zero-point, the unit length, or the angular orientation. This is reflected in the fact that all purely geometric concepts, like "being parallel to" etc., are not changed under certain transformations of the coordinate system. But such transformations are nothing but certain 1:1-permutations of the set points of the Euclidean space into itself. One important group are the *isometric transformations*, which corresponds to movements of rigid bodies - translation and rotation of the coordinate system. A more extensive class are the *similarity transformations* - they do not preserve distance, as isometric transformations, but may increase or decrease the distance uniformly in all directions - changes of unit length. Klein defines the purely geometric concepts in Euclidean geometry as those which are invariant under all similarity transformations. They preserve the *ratio* of two distances. For instance, "being a straight line", "being parallel to", "forming such and such angle" are purely geometric concepts in this sense, but not "being one meter long" etc.

If the class of transformations under which concepts have to be invariant is generalized, one obtains weaker and more abstract geometries. In *affine geometry* we consider concepts which are invariant under *affine transformations*: here, also ratios of distances may change, but colinearity and between-ness are preserved, i.e.

a line remains a line. Every triangle will be transformed into a triangle, but the angles are not necessarily preserved. Finally in *topology*, one considers only those concepts which are preserved by *continuous transformations*, a still wider class of transformations - they may bend lines into curves, but never break apart closed curves. Here a triangle may be transformed, e.g., in a circle, but never into, say, two triangles.

Recall that every such transformation is a permutation of the set of points, i.e. the objects of the domain. The natural idea of Tarski (1986) is that if we go on in the process of abstraction and consider only those concepts which remain invariant under *all* permutations of the domain, then this maximal generalization should lead us exactly to the logical concepts terms.

The invariance condition can not only be applied to *terms*, but also directly to set-theoretical *entities*: a set-theoretical entity constructed from a domain Dm is *logical* iff it remains invariant under arbitrary permutations of Dm. A term, then, is logical iff its extension is a logical one. Tarski (1986) lists which set-theoretical entities constructible from a given domain Dm are logical: (1) *Never* an individual. (2) Among the classes (subsets of Dm), only the universal and the empty class. (3) Among the relations (subsets of $Dm \times Dm$) we have four: the empty class, $Dm \times Dm$, the diagonal of $Dm \times Dm$, and the complement of this diagonal. This reflects nicely what we can express by logical symbols: $Fx \wedge \neg Fx$ or $Rxy \wedge \neg Rxy$ as the empty class, $Fx \vee \neg Fx$ or $Rxy \vee \neg Rxy$ as the full Dm or $Dm \times Dm$, $x=y$ (the identity relation) as the diagonal of $Dm \times Dm$, and $\neg x=y$ as the complement of this diagonal. (3) If we turn to classes of classes, i.e. to higher order extensional properties, things gets more complicated. But it turns out, as Tarski concludes, that a higher order class is logical iff it depends solely on the cardinality of its argument classes. If we turn to relations between classes, then the logical relations are, e.g., the standard set-theoretical relations of inclusion, disjointness, overlapping, etc. A similar thing holds if we view for logical functions.

Also quantifiers can be viewed as higher order predicates. They are logical if their extension depends only on the cardinalities of their argument classes, or on set-theoretic relations between their argument classes and the domain. Quantifiers are intensively discussed by Gila Sher (1991). A simple possibility of defining the extensions of quantifiers discussed by her (1991, pp. 11ff) to view their extensions as *partial* functions, defined for each chosen domain. If the quantifier is applied to a formula A with only one free variable x, then $I(\forall x)$ is the function $I(\forall x): Pow(Dm) \to \{t,f\}$ such that $I(\forall x A) = t$ iff $Dm = I(A)$ - or, alternatively in *free logic*, iff $Dm \subseteq I(A)$, else $I(\forall x A) = f$. Similarly, $I(\exists x A) = t$ iff $I(A) \not= \emptyset$ - or, alternatively in *free logic*, iff $Dm \cup I(A) \not= \emptyset$, else $I(\exists x A) = f$. In the general case, one has to state these definitions, of course, recursively, which is more complicated and is omitted here. Obviously, the extension of quantifiers defined in this way satisfies the invariance condition; thus the standard quantifiers are logical terms according to (IV). Gila Sher then turns to Mostowski quantifiers - cardinality quantifiers like "There are exactly n x such that:", or quantifiers like "Most x are such that:", meaning that the cardinality of things such that ... is greater than the cardinality of the complement class (1991, pp. 14ff). Also their extensions can be specified by similar clauses; they depends only on cardinalities and thus they are logical quantifiers according to the invariance criterion. This was the reason given by Mostowski why his generalized quantifiers should be treated as genuine logical quantifiers.

What does the criterion (IV) imply for propositional connectives? If they are viewed as functions between truth-values, then they are trivially invariant under domain permutations, because domain permutations do not affect truth values. We get a more nontrivial picture if we define the extension of a sentence as the set of objects which satisfy it, as in Tarski's paper on the concept of truth (1933). Then sentences have only two possible extensions: the extension of a true sentence is the entire domain, and the extension of a false sentence is the empty set. These concepts are permutation-invariant, and thus also truth functions are permutation-invariant.

The invariance criterion is certainly a *necessary condition* for logicalness of terms. Let us take a breath and ask: is it also *sufficient?*

8. McCarthy's Puzzle (Round 4 of the T-C-Game)

Unfortunately, the answer seems to be *no*. As McCarthy (1981) has shown, it is possible to describe entities which are logical according to the invariance criterion also with help of *empirical* predicates or statements. As simple example of McCarthy is the following definition of a new propositional connective $---$:

(Def $---$): $p --- q$ is true iff $p\wedge q$ is true and snow is white, or $p \vee q$ is true and snow is not white.

Since the meaning of the metalanguage terms is, of course, not varied but let constant, the truth-function of the connective $---$ is *in fact* that of the conjunction, because snow is in fact white. So the extension of the connective $---$ is a constant truth-function, and is invariant under permutations of the domain. Yet $--$ is intuitively a nonlogical connective. For its extension does not determined by purely formal and language-internal conditions, but by empirical conditions - the determination of the extension of $---$ depends on the real world. In other word, the above semantical rule for \$---\$ is not a *language-internal* rule according to our previous distinction.

An more nontrivial example of permutation-invariant extensions expressed by nonlogical terms are contingent cardinality quantifiers like:

(3) There are as many x as there are planets, such that:
(as opposed to: There are 7 x such that:)

(cf. Sher 1991, pp. 64ff). The extension of a contingent cardinality quantifier is again permutation-invariant, but, of course, this extension is not determined by language-internal rules, but depends on the real world. It follows that the invariance criterion (IV) is *not sufficient* as a criterion for logical terms. In cannot be so. For if it were, then a sentence like "there are as many natural numbers lower-or-equal 7 as there are planets" would have to be counted as a *constant* sentence and thus would have to be **L**-determined (according to the transparency condition T) - but this sentence is certainly contingent. We have to go back to the modified Carnap criterion (C^+) with its reference to language-internal rules. Only this reference can give us the intended distinction between "genuine logical" versus "contingent" descriptions of permutation-invariant extensions.

On the other hand, the invariance criterion (IV) is still useful, namely as a necessary but insufficient *test criterion* for extensions which are supposed to follow from purely language-internal rules alone. This is also McCarthy's conclusion in (1981), although he sets up a much more difficult criterion than (C^+) (so far I am not clear about the relation of his suggestions to mine).

Note that the same objection can be made for "purely geometrical concepts". E.g., if we define the contingent concept "the angle between two empirically given lines a and b", then this concept will also be invariant under similarity transformations, but it cannot be called a purely geometrical concept, because in order to understand it we need contact to the real world. In contrast, "the angle of 45 degrees" is purely geometrical in the strict sense, understandable without perceiving particular objects in the real world. Thus we may apply the modified Carnap criterion also to geometry, as an alternative to Klein's criterion. If we do this, we get the following definition: a concept is purely Euclidean-geometrical iff its the extension is solely determined by the axioms of Euclidean geometry. Indeed, this is true for the extension of "the angle of 45 degree", but not for the extension of "the angle given by two empirically given lines a and b".

9. The Puzzle of Quantifiers - a Solution

The puzzle of quantifiers has still to be solved. Tarski's implicit transparency requirement (T) is violated in the standard conception of models and quantifiers. Gila Sher concludes that, in contrast to Tarski's view, (T) should given up (1991, pp. 45f). But (T) is central for 'genuine' logics in the Tarski-Carnapian spirit. If the interpretation of quantifiers should be coherent with criterion (T), then quantifiers must contain a *variable* component which refers to the variable domain. I want to suggest a analysis of quantifiers which is more refined than that of Etchemendy (1990, pp. 65ff; he views quantifiers as pairs: "every - thing"). A (standard) quantifier is a compound of three components:

(4) $\forall x_i$: the logical term $\forall$ + the variable term x + the auxiliary index $_i$; and similarly for $\exists x_i$.

x is the *bound variable*, and its interpretation is the domain: $I(x) = Dm$. To keep coherent we have to select one *unique* symbol of the language for the bound variable. If we would take more than one symbol, say $x, y, \ldots$, then we could assign to these bound variables different interpretations (according to the independence criterion (I)), and sentences like $\forall x Fx \leftrightarrow \forall y Fy$ would no longer be logical theorems; we would rather have the situation of a many-sorted predicate logic. The auxiliary index $_i$ becomes necessary because we use one unique symbol for the bound variable. It indicates the argument place in the scope of a quantifier to which the quantifier refers, for instance, in sentences like $\forall x_1 \exists x_2(Fx_1 \to Rx_1x_2)$. Concerning free variables we have the choice of either choosing only variables different from x_i for them, or letting the context determine whether a variable x_i is bound or free.

The symbols $\forall$ and $\exists$ are logical symbols, and their fixed extensions are certain functions on sets. Let SET be the class of all sets of *possible* (i.e. conceivable, constructible) objects. Then the fixed extension of $\forall$, if applied to a formula A with just one free variable x, is the function $I(\forall): SET \times SET \to \{t, f\}$ such that $I(\forall)(I(x),I(A)) = t$ iff $I(A) = I(x)$, else $= f$. Note that $I(x)$ is the domain Dm and $I(A)$ is the extension of the unary "complex" predicate A. In the general case, the definition has again to be stated recursively, which is more complicated and omitted.

Given this analysis of quantifiers, then semantics of classical first order logic turns out to be not "genuinely logical". For, it contains the *cross term restriction* that the extensions of individual terms and predicates have always to be taken from $I(x)$ ($I(a) \in I(x)$ and $I(R^n) \in I(x)^n$). On this reason, the definition of the extension of the classical $\exists$-operator reduces to: $I(\exists): SET \times SET \to \{t, f\}$ such that $I(\exists)(I(x),I(A)) = t$ iff $I(A) \not= \emptyset$, else $= f$. To obtain a genuine first order logic satisfying (I), we must drop this cross term restriction and admit that also objects which are not in the domain I(x) may be taken as extensions of individual terms or predicates. The definition of the extension of a genuinely logical existential operator must be: $I(\exists): SET \times SET \to \{t, f\}$ such that $I(\exists)(I(x),I(A)) = t$ iff $I(x) \cap I(A) \not= \emptyset$, else $= f$. This implies that the classical theorem $Fa \to \exists x Fx$ is no longer a 'genuine' logical theorem, but a meaning postulate. In other words, it follows that *the genuine first order logic* is *not classical but free logic*. I think that this is not a drawback but, in the contrary, a merit of this account. It was repeatedly emphasized by philosophers that the question whether every entity which is subject of a predication must also exist goes *beyond* logic.

In the treatment of quantifiers as terms with the variable component x we had still to introduce one fixed and "hidden" semantical parameter, namely the *universe of all possible objects* (and the class of all sets of them). This universe is the class of all 'entities' which may figure as objects of possible interpretations. Hence, it is a universe of *possible* objects, of purely *conceptual* objects. Of course, it is difficult to describe how this universe may get fixed; we need not answer this extremely difficult question here. Etchemendy points out that certain 1st order sentences are only consistent if this universe of possible objects in infinite. E.g., the 1st order sentence which asserts that the relation R is irreflexive and transitive and has no greatest element - this sentence can be true only if the domain is infinite, and hence only if the universe of possible objects is infinite. Etchemendy concludes from this argument that the logical truth of 1st order statements may depend "on the actual seize of the universe" (1990, p. 118). But the phrase "actual seize of the universe" is heavily misleading: it suggerates this seize to be a matter of extralogical, contingent facts. But the universe is that of all possible objects – far away from any actuality. Being aware of this we need not be astonished that certain properties of this universe are relevant for logical truths. Let us make this more precise.

The *semantical* view of logic – the view that logic is not only matter of syntactic calculi but has a semantic and conceptual foundation – can be expressed as follows:

(S) A sentence is **L**-possible (**L**-consistent) iff there exists a possible extensional interpretation which makes it true.

This very view implies *straightforwardly* that the notion of **L**-truth will depend on certain properties of the universe of all possible objects, for instance, on its cardinality. For, in the presence of (S), the sentence "It is **L**-possible that the domain is infinite" is metalogically equivalent with the sentence "The universe of possible objects is infinite". Hence, the semantic conception of logic *must* imply that logical truths will to some extent depend on properties of this conceptual universe. If we consider *set-theory* as the theory of the conceptual universe, then logical truths will be dependent on certain axioms of set theory. But these set-theoretic axioms are not at all statements about extralogical, contingent matters, but about conceptual matters - matters of *metalogic*. Gila Sher as well as Tarski discuss the question whether the discussed semantical criteria of logicalness - be it (C$^+$) or (I) - will not automatically turn *all mathematical concepts* into logical concepts. I will not try to answer this extremely difficult question here; but it is obvious that at least many set-theoretical concepts become logical concepts in our view- again a welcome consequence.

10. Consequences for Modal Logic

Modal logics are not genuine logics, but systems of extralogical analytical postulates or rules about the intensional sentence operator $\Box$. (It should be clear that this classification of modal logics as not genuine logics but as systems of analytical principles *does in no way diminish their philosophical importance* - but it avoids several confusions about the question "what is the right modal logic?". The hidden semantical parameter in modal logics is the entire Kripke frame $\langle W,R \rangle$. We have to view this frame as the variable extensional interpretation of the modal operator $\Box$: $I(\Box$ = $\langle W,R \rangle$. Modal logics are sets of modal formulas which are true for certain classes of frames from which the interpretations of $\Box$ are taken (e.g. all universal frames, which gives S5, etc.). The metalogical principles characterizing certain frame classes are extralogical meaning postulates.

How would a *genuine modal logic* look like? It should have a *fixed frame* $\langle W,R \rangle$. Naturally, W should be the set of *all* logically possible words, and R be the universal relation on W. Indeed - this is nothing but Carnap's *original conception of modal logic* (1947, pp. 173ff; and 1946, system MFL). It is an historical error to think that Carnap's modal logic was S5. Only in the propositional part of his paper (1946, system MPL) Carnap deviates from his original idea and introduces closure under substitution to arrive at a system equivalent with the Lewis system S5. But modal logic according to his original idea is much stronger than S5. In the genuine Carnapian modal logic it holds that $\Box A$ is logically true if and only if A itself is logically true, for arbitrary formulas A (1947, p. 174; convention 39-1). Thus, e.g., for every atomic variable p, $\Diamond p$ is logically true and $\Box p$ is logically false - moreover, every completely modalized sentence will be **L**-determined. Carnap's genuine modal logic has very unusual properties - for instance, it is *not closed under substitution for propositional variables*, and its rules are *nonmonotonic*.

References

Rudolf Carnap (1928): *Der Logische Aufbau der Welt,* Felix Meiner, Hamburg.

Rudolf Carnap (1937): *The Logical Syntax of Language*, Routledge & Kegan Paul, London (German Original 1934).

Rudolf Carnap (1946): "Modality and Quantification", in: *Journal of Symbolic Logic* 11, pp. 33 - 64.

Rudolf Carnap (1947): *Meaning and Necessity*, Univ. of Chicago Press, Chicago.

John Etchemendy (1990): *The Concept of Logical Consequence*, Harvard University Press, Cambridge/Mass.

Kurt Gödel (1931): "Über formal unentscheidbare Säze der Principia Mathematica und verwandter Systeme I", *Monatshefte für Mathemathik und Physik* 37, pp. 349-360.

Fritz Mautner (1946): "An Extension of Klein's Erlanger Program: Logic as Invariant Theory", *Amer. J. Math.* 68, pp. 345 - 384.

Alfred Mostowski (1957): "On a Generalization of Quantifiers", in: *Fundamenta Mathematicae* 44, pp. 12 - 36.

Timothy McCarthy (1981): "The Idea of a Logical Constant", in: *Journal of Philosophy* 78, pp. 499 - 523.

Gerhard Schurz (1994); "Logical Truth: Comments on Etchmendy's Critique of Tarski", in: Twardowski, B./Wolenski, J. (eds.), *Sixty Years of Tarski's Definition of Truth*, Philed, Kraków, pp. 78 -95.

Gerhard Schurz (1997): *The Is-Ought Problem. An Investigation in Philosophical Logic*, Trends in Logic Vol 1, Kluwer, Dordrecht.

Gila Sher (1991): *The Bounds of Logic*, MIT Pres, Cambridge/Mass.

Alfred Tarski (1933): "The Concept of Truth in Formalized Languages", reprinted in: Tarski (1956), pp. 152 - 278.

Alfred Tarski and A. Lindenbaum (1934/35): "On the Limitations of the Means of Expression of Deductive Theories", reprinted in Tarski (1956), pp. 384 - 392.

Alfred Tarski (1936): "Über den Begriff der logischen Folgerung", in: *Actes du Congrés International de Philosophie Scientifique* 7, pp. 1-11. Translated into English as "On the Concept of Logical Consequence", in: Tarski (1956), pp. 409 - 420; page references refer to this English reprint.

Alfred Tarski (1956), *Logics, Semantics, Metamathematics*, Clarendon Press, Oxford.

Alfred Tarski (1986): "What are Logical Notions", in: *History and Philosophy of Logic* 7, pp. 143-154 (manuscript of a 1966 Tarski lecture edited by John Corcoran).