

Journal of Philosophy, Inc.

A Note on Extension, Intension, and Truth

Author(s): Anil Gupta and Nuel Belnap

Source: *The Journal of Philosophy*, Vol. 84, No. 3 (Mar., 1987), pp. 168-174

Published by: Journal of Philosophy, Inc.

Stable URL: <http://www.jstor.org/stable/2026597>

Accessed: 28/05/2009 11:06

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Please contact the publisher regarding any further use of this work. Publisher contact information may be obtained at <http://www.jstor.org/action/showPublisher?publisherCode=jphil>.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

JSTOR is a not-for-profit organization founded in 1995 to build trusted digital archives for scholarship. We work with the scholarly community to preserve their work and the materials they rely upon, and to build a common research platform that promotes the discovery and use of these resources. For more information about JSTOR, please contact support@jstor.org.



Journal of Philosophy, Inc. is collaborating with JSTOR to digitize, preserve and extend access to *The Journal of Philosophy*.

<http://www.jstor.org>

A NOTE ON EXTENSION, INTENSION, AND TRUTH*

IT is common knowledge that two predicates may coincide in extension but differ in intension and that, for any predicate, one can construct an infinity of coextensional predicates that differ in intension. Hence the concept that a predicate expresses—and its intension—cannot be determined from the extension of the predicate alone. Something stronger also appears to be the case. If a predicate has the same extension as a given concept, then one cannot determine whether the predicate expresses the concept simply on the basis of extensional information. The purpose of this note is to show that this appearance is illusory. We will construct an example in which a predicate G belonging to a first-order language L has the same extension as a concept \mathcal{C} , and yet merely by knowing the extensions of the nonlogical constants of L one can determine that G does not express \mathcal{C} .

The example is based on the fact that under certain conditions—conditions under which certain kinds of vicious reference are absent—a classical language L can contain a predicate that is extensionally equivalent to its own truth. Such a predicate we shall call a *T-predicate for L* .

Here is one simple way of constructing a classical language that contains its own T-predicate. Begin with an ordinary first-order language with the usual logical resources ($=$, \sim , \supset , \forall , etc.), and add to it a one-place predicate G and quotational names for all the sentences. Let the resulting language be L . Take any model M of the G -free fragment of L that meets the following conditions: (i) the domain D of M contains all the sentences of L ; (ii) the quotational names receive their intended interpretation; (iii) the interpretations of the predicates and the nonquotational names do not distinguish among the sentences of L .¹ This last condition means that no names other than the quotational names pick out a sentence of L ; that if a one-place predicate is true of a sentence of L , it is true of every sentence of L ; and similarly for the n -place predicates. Now, if we assign to G an interpretation, say g , we obtain a standard classical model $M + g$

* We wish to thank Mark Wilson and Michael Kremer for their valuable comments on the first draft of this note. Kremer helped clarify the role of symmetry in the argument and suggested the appropriateness of the term 'supervenience'.

¹ What we are calling "models" are sometimes called "structures." The interpretation of a predicate in a model M determines and is determined by its extension in M . For concreteness, we suppose below that the interpretation of a one-place predicate in M is a function from the domain of M into the truth values *The True*, *The False*.

for the whole of L . Let L_{M+g} be the resulting *interpreted* language. We can determine by the usual Tarskian method the sentences of L_{M+g} that are true and those which are false. Define an operation τ_M on the possible interpretations of G in M in the following way: for any interpretation g , $\tau_M(g)$ is the interpretation that assigns the value *The True* to the sentences true in L_{M+g} and the value *The False* to the rest of the objects in D . Observe that G is a T -predicate for L_{M+g} if and only if g is a fixed point of τ_M [i.e., if and only if $g = \tau_M(g)$]. It can be shown that τ_M has a fixed point and indeed a unique one.² If g is this fixed point then L_{M+g} is a classical language that contains a T -predicate for itself.

The language L_{M+g} constructed above has no vicious reference of any kind. Presence of certain kinds of vicious reference destroys the possibility of the above construction. For example, suppose that the model M' is just like M except that in M' an individual constant b denotes the sentence $\sim Gb$. It can be shown that $\tau_{M'}$ has no fixed points. (The argument here parallels the proof of Tarski's indefinability theorem.) So there is no interpretation g that makes G a T -predicate for $L_{M'+g}$. On the other hand, presence of certain other kinds of vicious reference leaves intact the construction given above. It is this that we exploit in our example.

Suppose that L contains, in addition to the other constants, a name c . Suppose that M^* is a model of the G -free fragment of L that meets the conditions (i) and (ii) given above and, apart from the constant c , meets condition (iii) also. Finally, suppose that the denotation of c in M^* is Gc . Now one can show that τ_{M^*} has exactly two fixed points. With one of them, say g , we obtain a classical language L_{M^*+g} in which Gc is true. With the other, say g' , we obtain the classical language $L_{M^*+g'}$ in which Gc is false.

Imagine a community C in which the language L is so used that its interpretation is, as a matter of fact, given by $M^* + g$. Imagine, in particular, that the predicate G in the language of this community means "is a sentence that the computer XYZ is programmed to

² We have assumed for simplicity that the language contains no function symbols. This assumption is dispensable at the cost of adding a fourth condition to the three stated above. Further, the conditions given can be considerably weakened. For more details on this and for a proof of the existence and uniqueness of the fixed point, consult Anil Gupta, "Truth and Paradox," *Journal of Philosophical Logic*, xi, 1 (February 1982): 1-60; also reprinted in Robert L. Martin, ed., *Recent Essays on Truth and the Liar Paradox* (New York: Oxford, 1984). The idea of fixed points, in the context of a three-valued approach to the theory of truth, was introduced most notably by Robert L. Martin and Peter Woodruff, "On Representing 'True-in- L ' in L ," *Philosophia*, v, 3 (July 1975): 213-217; and by Saul Kripke, "Outline of a Theory of Truth", this JOURNAL, LXXII, 19 (Nov. 6, 1975): 690-716. Both of these papers are reprinted in Martin's *Recent Essays*, *op. cit.*

produce,” and that, as a matter of fact, the computer is programmed to produce all (and only) the true sentences of L_{M^*+g} , including the sentence Gc . Since g is a fixed point of τ_{M^*} , the interpretation of G in this community’s language is given by g . So the predicate G as used by this community is coextensional with the following two (intensionally nonequivalent) concepts:

- being a true sentence in the community’s language (\mathcal{C})
- being a sentence that the computer XYZ is programmed to produce (\mathcal{C}')

We will argue that G cannot mean \mathcal{C} . A linguist who discovers that the predicate G in the community’s language is coextensional with the concepts \mathcal{C} and \mathcal{C}' will have reasons to conclude, on the basis of the extension of G alone, that it does not express \mathcal{C} . The following terminology will prove useful: let us say that a predicate of L is a *truth-predicate* if and only if it means “true in L .” Our claim, then, is that, although G is a T-predicate for the community’s language, it cannot be its truth-predicate.

Let us observe first that the hypothesis that G expresses \mathcal{C}' has certain virtues over the hypothesis that G expresses \mathcal{C} . On the first hypothesis Gc says of itself that it is a sentence that the computer XYZ is programmed to produce, and it is true because the computer *is* programmed to produce Gc . There is no mystery why Gc is true. On the second hypothesis Gc says of itself that it is true and, thus, is like the Truth Teller in English: “This very sentence is true.” Now it is a mystery why Gc is true in the community’s language. We can—in some sense of ‘can’—say that the Truth Teller is true, and we can also say that it is not true. But there is little to choose between these options. Both are equally arbitrary. So, on the hypothesis that G expresses truth, we have the puzzle: how come the Truth Teller is true and not false in the community’s language? This observation motivates, we hope, our claim. It does not prove the claim. However, the hypothesis that G expresses \mathcal{C} conflicts with a basic feature of truth. We call this feature *the supervenience of the semantic interpretation of truth*. It is related to, but distinct from, Ramsey’s Redundancy Theory.

It has often been remarked that the sentence “‘snow is white’ is true” says nothing more nor less than the sentence ‘snow is white’; that ‘is true’ is eliminable from and redundant in the first sentence. Whether this observation can be extended to all the sentences of a language containing its own truth-predicate depends on the logical resources of the language. If the language has propositional variables and quantifiers, the redundancy thesis is probably true. But if such quantification is not available and the language has ordinary

first-order quantifiers, the redundancy claim cannot hold for all the occurrences of 'true'. Nevertheless, even here the following kind of redundancy does obtain: *to evaluate the sentences of a first-order language that contains its own truth-predicate one does not need the semantic interpretation of the truth-predicate*. For example, suppose we wish to determine the semantic status of a sentence such as

$$(1) \quad (\forall x)(Fx \supset Gx)$$

in a model. Suppose we are given the domain of the model and the interpretation of F . Suppose also that F is true of just one object in the domain: the sentence Hd . Now, to determine the semantic status of (1), we would need to know the interpretation of G in the model. This is so *except* when G is a truth-predicate. In this case the status of (1) is entirely determined by the interpretations of F , H , and d . More generally, the point is that, if a first-order language contains a truth-predicate G , then the status of all the sentences of the language is entirely determined by the semantic interpretation of the G -free part. The interpretation of G —whatever it be: two-valued or three-valued or via a rule of revision or some other³—supervenes upon the interpretation of the G -free part. This feature we are calling "the supervenience of the semantic interpretation of truth".⁴ In reference to the community C imagined a little earlier, this feature implies that if G is a truth-predicate then the interpretation of G is fixed by the interpretation of the G -free part, i.e., by the model M^* .

Let us now imagine the community C in a slightly different situation s' . (Let us call the situation considered earlier s .) Suppose that in s' the meanings of the various constants in C 's language are as in s ; in particular, G in s' still means "is a sentence that the computer XYZ is programmed to produce." Suppose that, in other respects also, s' is just like s except that in s' the computer XYZ is programmed to

³ Two-valued theories have been sketched by Charles Parsons and Tyler Burge; three-valued approaches were developed, as already indicated, by Kripke, Martin and Woodruff, and others; the revision approach was initiated by Gupta, Hans Herzberger, and Belnap. For these and other approaches to the theory of truth consult the papers reprinted in Martin's *Recent Essays*. The point made in this note holds, as far as we can tell, for all approaches.

⁴ The supervenience described above is a special case of a much more general feature of truth. Suppose that L_1, \dots, L_n are languages that contain truth-predicates for each other. (One extreme possibility here is that they form the bottom part of a Tarski hierarchy. At another extreme each may contain truth-predicates for all the others.) Now, to determine the semantic status of the sentences of L_1, \dots, L_n , one does not need the interpretation of the truth-predicates. These are redundant once the interpretations of the other constants of L_1, \dots, L_n are fixed.

Note that supervenience does not hold, in general, of T-predicates, but only of truth-predicates.

produce the sentence $\sim Gc$ and all the rest of the truths of $L_{M^*+g'}$. As before, since g' is a fixed point of τ_{M^*} , G is coextensional with the concepts \mathcal{C} and \mathcal{C}' . Now observe that the hypothesis that G is a truth-predicate in s' is no less (or more) plausible than the hypothesis that it is a truth-predicate in s . The two cases are exactly parallel. The only difference between them is that with the first hypothesis one is led to say that the Truth Teller is true and, with the other, that the Truth Teller is false. Neither of these consequences is any worse than the other. However, the supervenience feature discussed above implies that one cannot accept both of the hypotheses. For the G -free part of the language has the same interpretation in the two situations: its interpretation is fixed by M^* . So our hypothetical linguist would know, simply on the basis of the extension of G , that both of the hypotheses cannot be true. Symmetry of the two hypotheses leads to the conclusion that in neither situation is G a truth-predicate.

There are two features of the above argument that will be of concern even to a sympathetic reader. First, the claim of symmetry of the two hypotheses will appear debatable. In fact, some philosophers maintain that the Truth Teller is simply false. (The view that it is simply true is not found in the literature, as far as we know.) These philosophers will not regard the two hypotheses as equally plausible. They would favor the latter hypothesis over the former. Secondly, in the argument given we have not carefully distinguished between what is the case in a situation from what the hypothetical linguist knows to be the case. It may appear that the linguist needs to know more than we attribute to him, if he is to arrive at the desired conclusion.

Concerning the first point: Notice that our original claim—that the linguist will be able to conclude in the first situation s that G is not a truth-predicate—must be granted by anyone who denies symmetry and holds that the Truth Teller is false. Supervenience now implies that, if the interpretation of the G -free part of the language is given by M^* and G is a truth-predicate, then the interpretation of G must be g' . So the linguist will be able to conclude that although G is a T-predicate, it is not a truth-predicate. For the interpretation of G in s is g , and, had G been a truth-predicate, its interpretation would have been g' .

By similar reasoning, if the Truth Teller is held true, then s' yields the desired example. Hence, even if symmetry be denied, it must be admitted that, in one of s and s' , G agrees in extension with truth although it is not a truth-predicate. The role of symmetry in our argument is minimal. It entitles us to say that, in *both* s and s' , G fails

to be a truth-predicate. It is, however, not needed in order to show the existence of the desired example.

Concerning the second point: We need to reconstruct the argument from the viewpoint of the linguist—call her Sally. Recall that Sally knows the extension of G in s and she knows that G is coextensional with the concept \mathcal{C} . This enables her to conclude that the interpretation (in s) of the G -free part of the language is given by a model M^* (she may not know which one) and that τ_{M^*} has at least one fixed point (for she knows that G is coextensional with the truth-concept \mathcal{C}). Since the language has quotational names and identity, she knows by examining the extension of \mathcal{C} that Gc is self-referential and that this is the only kind of vicious reference in the language. [She needs to know that the quotational names have their standard interpretation and that the predicates meet condition (iii) stated above. Since this does not require any intensional information about the language, we may assume that she does know this.] So she concludes that τ_{M^*} has exactly two fixed points g, g' , and that one of these, say g , gives the actual interpretation of G . (We do not need to assume that she knows anything about these fixed points other than that Gc is true in one and false in the other.) Believing in symmetry, she argues that, if it is possible that G be a truth-predicate in the community's language, then it is equally possible that G be a truth-predicate when the interpretation of the language is given by $M^* + g'$. So the hypothesis that G may be a truth-predicate in C 's language, she reasons, yields the existence of worlds w and w' such that in w (w') the interpretation of the whole of the language is given by $M^* + g$ ($M^* + g'$) and in both of the worlds G is a truth-predicate. But this violates the supervenience of truth. She concludes that G cannot be a truth-predicate. At the cost of greater tedium, this argument can be made more rigorous and precise. Let us hope the reader is satisfied with the above.

When a language contains its own truth-predicate, the presence of self-reference and cross-reference can result in pathological sentences such as the Truth Teller and the Liar ("This very sentence is not true") whose semantic status is clear neither to intuition nor to theory. It should be observed, however, that the argument given relies on no particular judgment on these troublesome sentences. The only property used is one that both intuition and theory agree on: the supervenience of the semantic interpretation of truth.

That a T-predicate for a language should sometimes fail to be its truth-predicate is not in the least surprising. For "T-predicate" is an extensional notion, whereas "truth-predicate" is an intensional one.

Any predicate that is coextensional with truth qualifies as a T-predicate, but, to be a truth-predicate, stronger conditions have to be met. What is curious about the examples given above is that sometimes one knows on the basis of extensions alone that a T-predicate does not express truth. In these examples a T-predicate is coextensional with truth only because it does not mean truth. If it *had* meant truth, its extension would *not* have been what it actually is.

Even within an extensional perspective, “T-predicate” and “truth-predicate” invoke different procedures, different points of view. To determine whether *G* is a T-predicate, we take the interpretation of *G as given* and using that interpretation we see whether *G* is coextensional with truth. If it is, then *G* is a T-predicate; otherwise it is not. But when we entertain the hypothesis that *G* is a truth-predicate, no interpretation of *G* is taken as given (because of the supervenience feature). We determine on the basis of the *G*-free part alone, the sentences that are true and those which are not true. Sometimes the two ways of proceeding give the same result. Sometimes, as in the examples constructed above, they give different results. It is this that enables us to infer from extensions alone that *G* is not a truth-predicate.

Examples of the sort given above can be constructed using other kinds of vicious reference, other kinds of logics, and other semantic concepts. In particular, examples of this sort can be given using the three-valued fixed points investigated by Saul Kripke, Robert Martin and Peter Woodruff, and others.

University of Illinois/Chicago

University of Pittsburgh

ANIL GUPTA

NUEL BELNAP