The Substitution Interpretation of the Quantifiers
Author(s): J. Michael Dunn and Nuel D. Belnap, Jr.
Source: *Noûs*, Vol. 2, No. 2 (May, 1968), pp. 177-185
Published by: Blackwell Publishing
Stable URL: http://www.jstor.org/stable/2214704
Accessed: 28/05/2009 15:06

The Substitution Interpretation of the Quantifiers

J. MICHAEL DUNN AND NUEL D. BELNAP, JR.
WAYNE STATE UNIVERSITY AND UNIVERSITY OF PITTSBURGH

I

Professor Ruth Barcan Marcus has suggested on several occasions a substitution interpretation of the quantifiers [1,2].[1] On this interpretation she reads

(1) $(\exists x)\varphi x$

as (2) Some substitution instance of φx is true.

And she reads

(3) $(x)\varphi x$

as (4) Every substitution instance of φx is true.

Professor Marcus believes that this reading dissolves the ontological import of the ordinary readings of the quantifiers, which read (1) as

(5) There exists at least one *thing* which is φ

and (3) as

(6) Every*thing* is φ.

Thus, to use her example, one might believe that

---

[1] The basic formal ideas of this interpretation were in a certain sense anticipated by E. W. Beth with his "reduced logic" in [3]. We are grateful to Professor Bas C. van Fraassen for bringing this to our attention.

The idea has also been employed by P. T. Geach in his "Quantification theory and the problem of identifying objects of reference" in *Acta Philosophica Fennica*, 1963, and by Henry S. Leonard in his "Essences, attributes, and predicates" in the *Proceedings of the American Philisophical Association*, 1963-64. Further, H. Leblanc has studied the interpretation in a pair of forthcoming papers, "A simplified strong completeness proof for QC=," and "A simplified account of validity and implication for quantification logic."

(7) Pegasus is a winged horse

is true (presumably on grounds like "fifty million mythologists can't be wrong") and still not believe that

(8) There exists at least one thing which is a winged horse

is true. And yet on the ordinary reading of the quantifiers, (8) would merely be a reading for

(9) $(\exists x)$ x is a winged horse,

and (9) would follow from (7) by the rule known (prejudiciously on her view) as *existential* generalization. However, Professor Marcus claims on her substitution interpretation of the quantifiers, (9) is a trivial and quite harmless consequence of (7), for it is to be read as

(10) Some substitution instance of 'x is a winged horse' is true,

and (7) is such a substitution instance.

Several years ago we discovered an odd, and seemingly undesirable, technical result about this substitution interpretation. Since there seems to be a widespread interest in using the substitution interpretation to handle certain problems involving opaque and non-designating occurrences of singular terms (both uses to which Professor Marcus has adverted), we thought it appropriate to make public this technical failing.[2] Let us quickly add that the interpretation which we shall first consider will not be the only interpretation that would count as a *substitution* interpretation. Indeed, Professor Marcus's own sketch of her interpretation in [2] differs in detail. But we ask the reader's indulgence until the point has been made about the precise substitution interpretation which we are about to give, for we shall then discuss how the same point may be extended to similar interpretations, and not extended to certain dissimilar interpretations. Finally, we shall do an about face, and defend the substitution interpretation against a number of potential misapprehensions.

---

[2] Essentially the same technical failing has been noticed independently by Professor Richmond H. Thomason in his doctoral dissertation, *Studies in the formal logic of quantification*, Yale University, 1965.

II

We begin by defining an *elementary language* as usual, with at least one predicate letter, a denumerable list of individual variables, perhaps some function letters, perhaps some individual constants, and the usual logical connectives and quantifiers. Terms and wffs will be as usual. By a *closed* term or wff we mean one with no free occurrences of individual variables (we shall call closed terms *names* and closed wffs *sentences*).

We now define the notion of a (*substitution*) *interpretation* for such a language. An *interpretation I* is a mapping of the atomic sentences into the truth values **T** or **F**. The *valuation* v_I determined by the interpretation is a mapping of the sentences into **T** or **F** that satisfies the following conditions:

1. if A is an atomic sentence, $v_I(A) = \mathrm{I}(A)$;
2. if $A = \overline{B}$, $v_I(A) = \mathbf{T}$ iff $v_I(B) = \mathbf{F}$;
3. if $A = B \,\&\, C$, $v_I(A) = \mathbf{T}$ iff $v_I(B) = \mathbf{T}$ and $v_I(C) = \mathbf{T}$;
4. if $A = (x)B(x)$, $v_I(A) = \mathbf{T}$ iff $v_I(B(t)) = \mathbf{T}$ for all names t.

We can now define as usual a sentence to be *valid* iff it is true in all interpretations, and as usual define a sentence A to be a *logical consequence* of a set of sentences X iff every interpretation in which every sentence in X is true is an interpretation in which A is true. Notice that we have defined these semantic notions of *truth*, *validity* and *logical consequence* only for sentences (the closed wffs of the language). We could have modified our definitions so as to apply to open wffs as well, but this would have been merely a distracting complication.

Notice that our definition of *validity* only makes good sense when the language has a denumerable number of names (since otherwise sentences involving quantifiers would be semantically equivalent to truth functions of atomic sentences). So henceforth we assume that the language under discussion is such. Under this assumption it is easily seen that in any of the standard formulations of the predicate calculus, a sentence is valid iff it is provable. This may be seen by examining any standard "tree" or "tableaux" proof of the completeness (under ordinary interpretations) of the predicate calculus, such as those of Beth, Schutte, Hintikka, or Kanger, where a non-provable sentence is falsified by a model in a domain consisting of the individual variables of the language. A judicious replacing of these variables by names gives the result for the substitution interpretation.[3]

But there is a more general completeness property, which Gödel showed to hold of a standard formulation of the predicate calculus given the ordinary (not the substitution) notion of *logical consequence*, namely, so-called *strong completeness*:

> SC: A sentence A is a logical consequence of a set of sentences X iff A is deducible from X.

It is easy to show that SC does not hold of any postulational formulation of the predicate calculus (where by a *postulational formulation* is meant a formulation where certain wffs are taken as postulates and rules of immediate inference are set down, and then a *deduction* of a wff A from a set of wffs X is defined as a finite sequence of wffs, each of which is either a postulate, or a member of X, or follows from preceding members of the sequence by one of the immediate rules of inference). For let X be the set consisting of the sentences $F(t_1), F(t_2), \ldots$, where $t_1, t_2, \ldots$ are all the names in the language, and let A be $(x)F(x)$. Clearly then A is a logical consequence of X, but it cannot be deducible from X. For if A were deducible from X it would be deducible from a finite subset of X (by the definition of *deduction*). But then the "if" part of SC would require that A be a logical consequence of this finite subset of X. But obviously, choosing X and A as we did, A is not a logical consequence of any finite subset of X.

Having noticed that SC fails under the substitution interpretation, one might hope for something weaker, perhaps that for every set of sentences X there is (maybe a different) set of sentences X' such that the sentences that are logical consequences of X are precisely the sentences that are deducible from X'. Thus, for example, given a standard formulation of the predicate calculus with the rules of *modus ponens* and generalization, we may form the set X' from the set X of the example above by adding to X the sentence $(x)F(x)$. This set X' clearly has precisely those sentences deducible from it which are logical consequences of the original set X.

---

[3] Professor van Fraassen has pointed out to us that this is essentially what Beth does in [3].

But this is too cheap, since such a trick will always work. Thus for any set of sentences X we can always take the set X' to be precisely the set of sentences that are logical consequences of X.

To make the situation more interesting, we put on the stipulation that both X and X' be recursive, as they happen to be in the example above. The question is then whether given any recursive set of sentences X, there is a recursive set of sentences X' such that for any sentence A, A is a logical consequence of X iff A is deducible from X'. Let us assume that we are considering an effective formulation of the predicate calculus so that the notion of a deduction of a sentence from a recursive set of sentences is recursive (perhaps because as usual the postulates of the predicate calculus are recursive, and the rules (finite in number) are recursive). Remembering that when the postulates of a theory are recursive, and the underlying predicate calculus is so effective, the theory is usually called *axiomatic*, we may view a recursive set of sentences X as serving as the postulates for an axiomatic theory. The question is then whether given an axiomatic theory with postulates X, there is an axiomatic theory with postulates X' such that a sentence A is a logical consequence of X iff A is deducible from X'. To play on words, we ask: are the logical consequences of the axioms of an axiomatic theory always axiomatizable?

We now answer this general question in the negative. Let us consider an elementary language with one binary predicate letter $=$, one unary function letter $'$, two binary function letters $+$ and $\cdot$, and one individual constant 0. This notation was chosen for its arithmetical associations, for we now form a set of sentences X as follows. An atomic sentence $s = t$ is called *correct* iff when we evaluate s and t according to the usual arithmetical interpretations of our notation, $s = t$ turns out to be an arithmetically true sentence. Otherwise an atomic sentence is called *incorrect*. We let X be the set of all correct atomic sentences and the negations of all incorrect atomic sentences.

It is easily seen, and easily verified by an inductive proof, that a sentence A is a logical consequence (with quantifiers given the substitution interpretation!) of X iff A is arithmetically true. But it easily follows from a sufficiently general form of Gödel's Incompleteness Theorem that there can be no axiomatic theory such that the sentences provable in the theory are precisely the arithmetically true sentences. And since the set X is obviously recursive, we have our desired counter-example.

This counter-example can be extended to elementary languages that differ in various ways from the arithmetical language above. For example, the counter-example can be extended to the full predicate calculus (with denumerably many individual constants, denumerably many predicate letters of every degree, and denumerably many function letters of every degree) by picking appropriate symbols to serve as the arithmetical ones above. The only problem then is that the additional individual constants and function letters introduce new names which may be used to falsify in the substitution interpretation. The solution then is to treat these additional names like 0 in the construction of the new set X. By various similar syntactical nitpickings one can extend the counter-example to any elementary language (with a denumerable number of names) that has sufficient expressive power for arithmetic, perhaps in virtue of some function letters being eliminated in virtue of predicate letters in the standard way. But these details do not seem now required. The point about the substitution interpretation has been made.

Let us now turn to explaining wherein the notion of a substitution interpretation sketched in this paper differs from that in Professor Marcus's [2]. First, it differs in that Professor Marcus seems to assume that an ordinary interpretation with its domain of things determines the truth values of the atomic wffs. This is obviously only an inessential difference since we could imagine that the interpretation function I which was defined above was determined in whole or in part by such an ordinary interpretation. Further, it might be remarked that it would seem that the truth values of atomic sentences must be at least sometimes not determined by an ordinary interpretation to make sense of the remarks by Professor Marcus in [1] about 'Pegasus is a winged horse' being true even though there is no such thing as Pegasus. Professor Marcus in [2] was concerned with different questions, arising in modal logic, so it is of course understandable that the way that she did things there might differ from her intent in [1].

The second way in which the notion of a substitution interpretation sketched in this paper perhaps differs from that in Professor Marcus's [2] is in defining the truth of a quantified sentence only in terms of the substitution instances that involve the substitution of a name (a closed term) for the variable of quantification, whereas it simply is not clear whether Professor Marcus intends this or not. Let us suppose for the moment that she intends to allow the substitution of any term, even a variable, to count as a substitution instance. It is easy to see that all we have to do to allow this is to change our mapping I so that it assigns truth values to all atomic wffs, not just to atomic sentences (closed wffs) as it does above. Further, it should be clear that the same points could be made about this kind of substitution interpretation. Indeed, if anything, this kind of interpretation would be in worse shape since we would not have to assume the elementary language in question to have a denumerable number of names. We could use variables in their place.

Let us close off our discussion of the technical failings of substitution interpretations by saying that there is a concept of *logical consequence* based on a substitution interpretation to which our remarks do not apply. For example, there is a notion of *logical consequence* which permits enrichment of the elementary language in question by the addition of new names, rather than holding the names fixed, as we did above. The idea is that for a given language L, a sentence A should be a *logical consequence* of a set of sentences X iff for every extension of the language L to a language L' obtained by adding new names to L, every interpretation of L' in which every sentence of X is true is an interpretation in which A is true. For example, $(x)F(x)$ would not then be a logical consequence of $F(t_1), F(t_2), \ldots$, even though $t_1, t_2, \ldots$ are all the names in the language, because we could always add a new name b to the language and let $F(b)$ be false. We think that this is a good notion of logical consequence, for it immediately ensures an intuitively desirable property, to wit, that relations of logical consequence be preserved upon extension of a language. Furthermore, strong completeness may be established for it merely by definitional modifications of the strong completeness proof for the ordinary domain-and-values notion of logical consequence of Henkin [4]. Indeed, any reader familiar with Henkin's proof, will recognize immediately how much of our "new-names" notion of logical consequence is derivative from Henkin, for the addition of new names was the profound trick which made his proof go through.

III

The tone of our paper so far has been critical of the substitution interpretation, in that we have demonstrated that certain technical properties which we think desirable fail for it unless one is careful with the notion of logical consequence. However, this is a technical point that could be given a different evaluation; for example, if one was doing number theory one might want a notion of logical consequence in which the arithmetical truths were the logical consequences of ones axioms, and one might be willing to sacrifice completeness for this. We now want to balance out the tone of our paper by attempting to correct what we feel are a number of potential misapprehensions about the substitution interpretation; we have heard each in conversation.

1. The substitution interpretation is unclear or imprecise. No. The semantics given herein are exactly as rigorous as ordinary domain-and-value semantics. Just different.

2. The substitution interpretation involves an illegitimate use of "etc." and thereby absurdly tries to reduce quantificational logic to propositional logic. No. There is an explicit reference in the semantics to an infinite totality (of names), and reduction to propositional logic is manifestly impossible.

3. The substitution interpretation is merely a special case of the ordinary one, making the set of names the domain and making variables take names as values. No. On the substitution interpretation there is no domain and variables do not "take values" at all.

4. The intuitive quantifiers used in the meta-language in giving the substitution interpretation to the formal quantifiers must themselves be taken in the ordinary domain-and-values sense, so that after all one is not freed completely from "ontological import"; one commits oneself at least to the existence of names. No. The intuitive quantifier *can* be so taken, but there is no reason not to give *it* a substitution interpretation as well.

5. The substitution interpretation makes quantification essentially a metalinguistic device and quantified statements "about" pieces of language. No. And this rejoinder is independent of the last, for it remains true even when the intuitive quantifiers used in giving the semantics are of the ordinary domain-and-values kind. It is of course true that the semantic description of the quantifiers is metalinguistic, but this applies with exactly the same force to the semantic description of, say, conjunction. This is because all semantic descriptions are metalinguistic. But if to describe the semantics of conjunction by saying that $A \& B$ is true if and only if both A and B are, does not make conjunction essentially metalinguistic and $A \& B$ "about" its conjuncts, then to describe the semantics of the universal quantifier by saying that $(x)A(x)$ is true if and only if all its substitution instances are, does not make uni-

versal quantification essentially metalinguistic and does not make $(x)A(x)$ "about" its instances.

6. The substitution interpretation is without utility. No. Putting aside the uses to which Professor Marcus adverts, it will often be somewhat easier to use whenever, were we to use the domain-and-values interpretation, every entity in the domain would be denoted by some name. And this happens often enough to be interesting, as in syntax, where we usually have a name for everything (i.e, every expression) we wish to talk about.

And speaking of syntax, we mention in closing a marvelous application of the substitution interpretation that comes about by observing that the substituends of variables need not be restricted to be names (nouns). Variables may occupy the place of expressions of other grammatical categories, for example, sentences. With this convention firmly in mind, we can say things like

(11) $(P)(Q)$ if "P" and "Q" are sentences, then "P & Q" is a sentence, and "P" and "Q" are its conjuncts.

And we can do so *without* running the risk of making a mistake about use and mention through inadvertently referring to the sixteenth letter of the English alphabet by using " "P" ".

Indeed, if we also have the convention that signs of the object language are used autonymously as names of themselves, then we may drop the quotes *without* having to introduce any quasi-quotes or other ugly notational conventions:

(12) $(P)(Q)$ if P and Q are sentences, then P & Q is a sentence, and P and Q are its conjuncts.

The substitution interpretation thus gives logicians a way of preaching what they practice.[4]

References

1. Ruth Barcan Marcus, "Interpreting quantification," *Inquiry,* V (1962): 252-259.
2. Ruth Barcan Marcus, "Modalities and intensional languages," *Boston studies in the philosophy of science,* ed. Marx W. Wartofsky (Dordrecht, Holland, 1963): 77-96.
3. E. W. Beth, *The foundation of mathematics* (Amsterdam, 1965), Section 89: 263-267.
4. Leon Henkin, "The completeness of the first order functional calculus," *Journal of symbolic logic,* XIV (1949): 159-166.

[4] This research was supported in part by National Science Foundation Grant GS-689.