

LOGICAL TRUTH AND INDETERMINACY

STEVEN E. BOËR

In various of his writings, W. V. O. Quine has attempted to characterize the conditions under which a sentence of a natural language counts as a logical truth. P. F. Strawson, in [6], lodges the objection that Quine cannot give an ultimately coherent account of logical truths in natural languages without recourse to the very sorts of intensional devices which he so steadfastly refuses to countenance. Since then, both Quine and his commentators have essayed replies to Strawson. My aim in this paper is to show that these replies, whatever their intrinsic merits, fail to supply *Quine* with an adequate answer to Strawson's original criticism.

1 To begin with, let us restate the Strawsonian objection in the context of Quine's most recent discussion of logical truth in [3]. Rather than talking directly about sentences of our natural language, Quine asks us to envisage a "regimentation" of that language into the idiom of standard first-order predicate calculus with identity, and he then proceeds to use the term 'sentence' in the sense of 'sentence of the regimented idiom.' He then offers three different but allegedly equivalent characterizations of logical truth (for a regimented language):

(L1) . . . a sentence is logically true if all sentences are true that share its logical structure. (p. 49)

(L2) A *logical truth*, then, is definable as *a sentence from which we get only truths when we substitute [open] sentences for its simple [open] sentences*. (p. 50)

(L3) A *logical truth* . . . is a truth thus obtainable [by substitution] from a valid logical schema. (p. 51)

Actually, (L3) itself could be expanded into three different definitions, depending on how one chooses to characterize the intermediary notion of a valid logical schema—e.g., in terms of the truth of all its substitution-instances, its being satisfied in all admissible models, or its provability via some demonstrably complete proof-procedure. But for our present purposes we may ignore this further complication.

Now (L2) and (L3) make explicit use of the notion of *substitution* in defining logical truth, and it is arguable that even (L1) does so covertly. It is here, at the juncture provided by the appeal to substitution, that Strawson attacks. Whether the operation of substitution is conceived of as applicable to sentences (of a regimented language) or to logical schemata, it is always demanded that the operation should be performed *uniformly*. Where a regimented idiom is concerned, uniform substitution is explained as putting the *same* open sentence for each occurrence of a given simple open sentence (or simple open sentence-schema) in the original sentence (or logical schema). But, Strawson asks, what does 'the same' mean here? With respect to a regimented language, 'the same open sentence' cannot merely mean *tokens of the same sentence-type*, i.e., typographically identical inscriptions. For many open sentences are ambiguous in virtue of residually containing certain ambiguous natural language predicates. For example, the sentence

- (1) $(x)(x \text{ is bright} \supset x \text{ is bright})$

comes from the logically true sentence

- (2) $(x)(x \text{ is round} \supset x \text{ is round})$

or from the valid logical schema

- (3) $(x)(Fx \supset Fx)$

by substitution which is "uniform" in the foregoing sense. Suppose, however, that the first occurrence of ' x is bright' in (1) has the sense of ' x is shiny' and that the second occurrence has the sense of ' x is intelligent.' Then (1) is surely false. By the lights of (L1) and (L2) it would follow that (2) is *not* logically true, since (2) appears to have a false substitution-instance in (1). And it would follow from (L3) that (1), although false, *is* a logical truth in virtue of its descent from (3).

Since both of the above results are clearly unacceptable, one's first instinct is to cast about for some extensional supplement to the bare requirement of typographical identity. The natural candidate is a requirement to the effect that the supplanting predicate (and, in cases where substitution is performed upon a sentence rather than a logical schema, the supplanted predicate as well) should have the same extension *at all its occurrences*. But, Strawson hastens to point out, this talk of predicates or open sentences must be talk either of types or of tokens. It cannot be the latter, since a token, *qua* scattered patch of ink, is not the sort of thing that could intelligibly be said to possess an extension at all. Nor can Quine opt for the former alternative, since one cannot talk of the extension of a predicate-type *at* an occurrence without recourse to intensionalism—without, that is, employing in one guise or another the notion of a predicate's being used in one of its *meanings*. Since the appeal to meanings is anathema to a right-thinking extensionalist, Strawson concludes that Quine must either lift his ban on intensional devices or else abandon all hope of giving an account of logical truth in natural languages or their regimented relatives.

2 How does Quine himself propose to deal with Strawson's objection? In [4], his only published remarks on the subject, Quine appears to grant Strawson a partial victory and concerns himself only with showing that the victory is not total. He avoids talk of logical truths in natural languages or their regimentations and contents himself with salvaging an extensional concept of logical truth

... for language regimentations in which this difficulty [regarding ambiguity] does not arise: *univocal* regimentations, in which the extension of a term stays the same from one occurrence to another. (p. 323)

The difficulty, of course, lies in saying what makes a regimentation "univocal" without importing any intensional devices. To solve this problem, Quine embarks on a circuitous route. Let T be any demonstrably complete proof-procedure for first-order predicate calculus with identity, modified so as to apply directly to sentences of a regimented language rather than to logical schemata. Then we may lay down the following definitions:

(D1) A language-regimentation is *univocal* iff every T -provable sentence of the regimentation is true.

(D2) A sentence of a univocal regimentation is *logically true* iff it is T -provable.

Clearly, if there is any extensionally ambiguous predicate lurking in the lexicon, it will show up sooner or later in a T -provable sentence which is false. Hence the exclusion of such sentences ensures the absence of such predicates.

Although Quine appears unruffled by the retreat from (L1)-(L3) to (D2), his commentators have been less acquiescent. In particular, it has been argued that, with respect to natural languages, one *can* speak of the extension of a predicate-type *at* an occurrence without recourse to intensionalism. Thus Gilbert Harman writes,

According to [Quine] a person associates varying paraphrases with ambiguous sentences. . . . The varying paraphrases represent what the speaker takes to be equivalent to the ambiguous sentence given a particular *context*. The notion of equivalence here is *not* that of "meaning equivalence" but rather the notion of an equivalence taken to follow fairly obviously from what the relevant group of people (possibly only the speaker or hearer) accepts. ([2], pp. 150-151)

Elaborating on this point, Marion Deckert writes,

The extensionalist need only take word types in relation to *contexts* to find something which is finely enough individuated to serve as the bearer of extensions. . . . The context of a sentence will be sufficient to determine the extensions of its general terms. Then it would always be possible to specify uniquely what the appropriate paraphrase would be by reference to the context of occurrence. . . . [A] sentence relative to a context would be found to be either logically true or not logically true by reference to its paraphrase in univocal notation. ([1], pp. 54-56)

There are three key notions involved here: "context," "paraphrase," and

“univocal notation.” We have lately explored Quine’s characterization of the last of these, and we must secure at least an equal understanding of the remaining two before our discussion can profitably proceed.

3 What is Quine’s conception of a *paraphrase*? It is clearly not the ordinary one. For in standard English one sentence is called a paraphrase of another just in case the former is a rewording of the latter which is held to be equivalent to it in *meaning*. Presumably, then, Quine thinks of paraphrase in terms of some weaker relation, palatable to an extensionalist. Since we are specifically concerned with the relation of paraphrase as it obtains between sentences of a natural language and those of univocal regimentations, we can do no better than to consult Quine’s own pronouncements concerning the aims and claims of regimentation. We find that by calling a regimented sentence *S'* a *paraphrase* of an ordinary sentence *S*, Quine means that

Its relation to *S* is just that the particular business that the speaker was on that occasion trying to get on with, with the help of *S* among other things, can be managed well enough to suit him by using *S'* instead of *S*. ([5], p. 160)

And we are cautioned that

... the speaker is the one to judge whether the substitution of *S'* for *S* in the present context will forward his present or evolving program of activity to his satisfaction. (*Ibid.*)

These passages also shed light on the elusive notion of *context*. Neither Harman nor Deckert strives to enlighten us as to the identity of those considerations which supposedly enable us to indicate an appropriate paraphrase for someone’s utterance. Surely, for example, a context is not just a group of sentences, i.e., the surrounding discourse. For although the verbal context in which a sentence is uttered will sometimes suffice to disambiguate it, and thus to indicate its appropriate paraphrase in univocal notation, this is obviously *not* the case on all (or even most) occasions of speech. Now Quine seems quite clearly to specify that the contextual features relevant to paraphrasing a sentence have primarily to do with the speaker’s *purposes* in uttering that sentence. To say, then, that paraphrase is relative to context is just to say that something counts as a paraphrase of someone’s utterance, not *in vacuo*, but only against the background of a particular project or activity which the speaker’s utterance was designed to further. If we let $A(M, S)$ be the particular activity in the course of which *M* utters *S*, and let *S'* be a sentence of a univocal regimentation of *M*’s language, then we may incorporate this notion of “context” into a precise definition of ‘paraphrase.’

(D3) *S'* is a *paraphrase* of *M*’s utterance *S* relative to $A(M, S)$ iff: *S'* is a sentence which, if it were intelligible to *M*, would be regarded by *M* as a pragmatically adequate substitute for *S* relative to $A(M, S)$.

The subjunctive clause concerning the intelligibility of *S'* to *M* is necessary

in light of the fact that M may never have learned any univocal regimentation of his native tongue; so we must concern ourselves with what M *would* do *if* he understood the sentence S' . And to say that M regards S' as a pragmatically adequate substitute for S (relative to $A(M, S)$) is just to say that M believes that, if he had uttered S' in place of S , he would have carried out $A(M, S)$ in at least an equally satisfactory manner.

Having already characterized logical truth in univocal regimentations, we must now see whether the notions of "context" and "paraphrase" (as explicated above) can be fitted together in such wise as to form a bridge between sentences of a natural language on the one hand and those of its univocal regimentation on the other. Letting \mathcal{L} and \mathcal{L}' be respectively a natural language and its univocal regimentation, we obtain from (D1)-(D3) the following definition:

(D4) A sentence S of \mathcal{L} is *logically true* relative to its utterance by M in $A(M, S)$ iff: every sentence S' of \mathcal{L}' which is a paraphrase of M 's utterance S in $A(M, S)$ is logically true.

It appears that (D4), or something relevantly like it, is the closest one can come on purely Quinean grounds to a definition of logical truth for natural languages. And (D4) undeniably *does* wear an air of plausibility. For, confronted with a palpably ambiguous sentence of a natural language and asked whether it is logically true, it seems quite natural to respond by inquiring into the character of the (actual or hypothetical) utterer, his purposes, the surrounding activity, etc., before venturing an answer.

4 The question I would now like to raise is whether Quine could accept (D4) without compromising any other views to which he is antecedently committed. The answer, I think, must be negative. For the apparatus of "paraphrase" which figures so prominently in (D4) is ultimately at odds with Quine's well-known thesis of the indeterminacy of translation. In outline, my argument is thus:

- (i) If Quine's general argument for the indeterminacy of translation is valid, then the translation of native expressions for assent and dissent is indeterminate.
- (ii) If paraphrase is determinate, then the translation of native expressions for assent and dissent is *not* indeterminate.
- (iii) Therefore, either Quine's argument for the indeterminacy of translation is not valid or paraphrase is not determinate.
- (iv) Definition (D4) is acceptable only if paraphrase is determinate.
- (v) Therefore, either Quine's argument for the indeterminacy of translation is not valid or (D4) is not acceptable.

Since the argument is truth-functionally valid, it remains only to establish the truth of (i), (ii), and (iv). Let us begin with (i).

Quine holds most translation to be indeterminate, not just in the weak (epistemic) sense that we can never know whether a particular translation is correct, but also in the strong (ontological) sense that there is *no fact of the matter* about correct translation in the absence of so-called "analytical

hypotheses," which outrun any possible evidence. The argument for strong indeterminacy is based on two premisses: first, that

All the objective data [the linguist] has to go on are the forces he sees impinging on the native's surfaces and the observable behavior, vocal and otherwise, of the native. ([5], p. 28)

and second that

...manuals for translating one language into another can be set up in divergent ways, all compatible with the totality of speech dispositions, yet incompatible with each other. (*Ibid.*, p. 27)

Whether these considerations really yield strong (ontological) indeterminacy is moot; but Quine thinks they do, and I shall not dispute the matter here.

The possibility of divergent translation-manuals arises from the possibility of a shift in analytical hypotheses, i.e., translational hypotheses for expressions shorter than sentences. Analytical hypotheses go beyond the behavioral-*cum*-environmental data; and since these data are the only source of evidence upon which the linguist can legitimately draw, analytical hypotheses transcend the realm of independent evidence altogether. Consequently, there can be no question of right or wrong about analytical hypotheses, nor any question of "supporting" one to the exclusion of another—except in the purely pragmatic sense that one such hypothesis might prove more congenial than another to our current purposes. It is the *need* for analytical hypotheses in achieving a unique translation that signals the advent of indeterminacy.

Within the framework of the indeterminacy thesis, Quine accords curious pride of place to the notions of *assent* and *dissent*. We are told, for example, that truth-function theory escapes the ravages of indeterminacy *because* the semantics of the truth-functional connectives is explicable in terms of assent and dissent, which in turn are supposedly identifiable by strictly behavioral criteria (cf. [5], p. 61). But Quine never tells us what these behavioral criteria are; instead, he seems to take for granted that native expressions for assent and dissent can be translated without reliance on analytical hypotheses. He even gives us a sketch of how this translation might be carried out in a particular case.

So we have the linguist asking 'Gavagai?' in each of various stimulatory situations, and noting each time whether the native assents, dissents, or neither. But how is he to recognize native assent and dissent when he sees or hears them? Gestures are not to be taken at face value; the Turks' are nearly the reverse of our own. What he must do is guess from observation and then see how well his guesses work. Thus suppose that in asking 'Gavagai?' and the like, in the conspicuous presence of rabbits and the like, he has elicited the responses 'Evet' and 'Yok' often enough to surmise that they correspond to 'Yes' and 'No', but has no notion which is which. Then he tries the experiment of echoing the native's own volunteered pronouncements. If thereby he pretty regularly elicits 'Evet' rather than 'Yok', he is encouraged to take 'Evet' as 'Yes'. Also he tries responding with 'Evet' and 'Yok' to the native's remarks; the one that is the more serene in its effect

is the better candidate for 'Yes'. However inconclusive these methods, they generate a working hypothesis. If extraordinary difficulties attend all his subsequent steps, the linguist may decide to discard that hypothesis and guess again. ([5], pp. 29-30)

Although this account of 'Evet' and 'Yok' sounds plausible, it does not square with the demands of the indeterminacy argument. For the translation of 'Evet' and 'Yok' can be seen to involve analytical hypotheses just as much as does the translation of 'Gavagai' and its kin. In the above passage Quine offers two tests for recognizing assent and dissent, viz., (a) questioning the native with 'Gavagai?' and the like; and (b) echoing the native's own pronouncements. Consider first (a). By what right does the *question-mark* appear in 'Gavagai'? Surely the identification of the interrogative mood indicator in the native's language requires analytical hypotheses, for there is no reason to believe that mood indicators are any more privileged with respect to behavioral scrutability than are the categorical copula and the identity sign, both of which are said to require analytical hypotheses for their translation.

If, on the other hand, we cannot identify the interrogative mood indicator in the native's language, then we have no ground for supposing that the native's responses to our utterances are *answers* to *questions*, and hence no ground for supposing that his responses have anything whatever to do with assent and dissent. For all we could ever know, our attempts at producing "questions" might succeed only in producing imperatives or optatives in the native's language, or even sentences of some totally alien mood not found in English. It might be thought that this tacit reliance of (a) upon analytical hypotheses is not too important, since one could always fall back upon (b) in translating native expressions for assent and dissent. So let us look more closely at (b).

On inspection, (b) appears to suffer from the same sort of defect as (a). If echoing the native's "pronouncements" is to elicit signs of assent or dissent, then it must be assumed that the echoed utterances were *statements*, i.e., declarative sentences of the native's language. For if we echo, say, an imperative or an interrogative, the native's response (if any) is not likely to have much to do with assent or dissent. And if the native's utterance was in some totally alien mood, there is no way of telling *what* his response might signify. Even the bare assumption that the native's language possesses only (or only some of) the moods found in English would amount to an analytical hypothesis, for it would 'impute our sense of linguistic analogy unverifiably to the native mind' ([5], p. 72). And the much more particular assumption that what we are engaged in mimicking is a declarative sentence, is clearly an analytical hypothesis. It seems, then, that if Quine is right about the sorts of consideration relevant to showing indeterminacy, then the translation of native expressions for assent and dissent is surely *indeterminate*.

Turning now to (iv), it is not difficult to see that the paraphrase relation must be determinate if (D4) is to do its job. For if paraphrasing, like most

translation, essentially involves analytical hypotheses and is thus indeterminate, then there will presumably be no fact of the matter about paraphrases, just as there is none regarding "correct" translations. And if there is no fact of the matter about paraphrases, then there is certainly no fact of the matter about logical truth in natural languages, since, according to (D4), the latter is to be thought of in terms of the former. But there is no point in trying to account for logical truths in natural languages unless it is supposed that there is some objective datum to be accounted for. So the acceptability of (D4) (as an account of a real phenomenon) is contingent upon the determinacy of paraphrase.

Now by way of establishing (ii), let us suppose that paraphrase is indeed determinate, so that we are dealing with an objective relation which obtains or fails to obtain whether or not anyone ever notices the fact. Then we may for convenience postulate a *paraphrase-machine*, i.e., an oracular "black box" that receives native sentences and produces a paraphrase in some univocal regimentation of the native's language. If the submitted sentence is declarative, then the machine produces an "interpreted" formula of first-order predicate calculus with identity. By calling the paraphrase an "interpreted" formula, I do not mean to suggest that it is a translation of the original sentence: I mean only that it will look like ' $(x)(\text{soltero}(x) \vee \text{gavagai}(x))$ ' rather than like ' $(x)(F(x) \vee G(x))$ '. The paraphrase-machine puts unfamiliar content into a familiar mold; it cannot, however, translate the residual general terms of the native's language. It is not clear what the machine would do to a non-declarative sentence, since the notion of first-order logic is designed primarily for application to declarative sentences. We might say that the machine would simply give some stock response like 'Not fully paraphrasable.' On the other hand, it is plausible to think that the machine might paraphrase as much of the submitted sentence as it could, putting in uninterpreted operators when its resources failed. Thus, for example, the interrogative sentence 'Is John fat?' might be paraphrased as ' $\#(\text{fat}(\text{John}))$ ', where '#' records the presence of an operator foreign to the canonical notation. Since nothing hinges on the outcome with non-declarative sentences, I shall adopt the latter view in what follows.

Since the intelligibility of the notion of a paraphrase-machine is guaranteed by the assumed determinacy of paraphrase itself, we may ask how such a machine would help the linguist translate native expressions for assent and dissent without the need for analytical hypotheses. The answer lies with the logical forms of the paraphrases. Any fully paraphrasable native sentence will be mapped onto a sentence of the univocal regimentation which, from the standpoint of its logical form, is either logically true, logically false, or contingent. If the paraphrase is logically true, then the linguist can echo the native's *original* pronouncement and be reasonably sure that the native's response is one of *assent*. (Repeating the paraphrase itself to the native might prove futile, since the native might not understand the univocal regimentation.) Similarly, if the paraphrase is logically false, then the linguist can echo the native's original utterance and be reasonably certain that the native's response is one of *dissent*.

The most important feature of this method is that it begs no questions about analytical hypotheses. In particular, no assumptions need be made about the *mood* of the native's original sentence: the paraphrase-machine settled this matter in advance. If the paraphrase is a well-formed formula in the canonical notation, without uninterpreted operators, then we know that we are dealing with a declarative sentence—or at least with a remark which “amounts” to a statement. Otherwise, we know that we are dealing with a non-declarative sentence—or at least with a remark with some force other than that of a statement. And since the tests envisaged above concern only declarative sentences, those tests require nothing not provided in the paraphrase itself.

Of course, we do not possess a paraphrase-machine, nor is it likely that anyone ever will. The story of the paraphrase-machine is merely a device to illustrate what *could* happen if paraphrase were determinate. In actual practice, the linguist would have to form his own hypotheses about paraphrases and test them against his observations of the natives' behavior. But the important point is that so long as paraphrase is assumed to be determinate, the linguist's procedure will be subject only to normal inductive uncertainty. And if the linguist can assert that it is *very probable* that such-and-such a sentence is a paraphrase of another, then he can use the method sketched above (minus the machine) to conclude that it is *very probable* that a given expression pertains to assent or dissent. Such a result is enough to show that the translation of native expressions for assent and dissent is *not* indeterminate if paraphrase *is* determinate.

5 Having established the truth of (i), (ii), and (iv), we may safely conclude that (v) is also true, i.e., that Quine cannot consistently accept (D4) *and* maintain that his general argument for the indeterminacy of translation is valid. Despite commentators' suggestions to the contrary, Quine *must* confine the notion of logical truth to univocal regimentations or else abandon his views regarding strong indeterminacy. This alternative is not happy, since abandoning the strong version of the indeterminacy of translation would have awkward consequences elsewhere. For it is the strong thesis of indeterminacy which is Quine's chief weapon in the ongoing struggle against intensionalism: it is what supposedly *eliminates* such creatures of darkness as meanings and propositions. Without the strong thesis, Quine would be forced to fall back upon the substantially weaker anti-intensionalist arguments based on allegation of “obscurity,” “lack of identity-criteria,” and so on. So, in the end, Strawson was right: Quine must either forgo an account of logical truths in natural languages or else materially weaken his anti-intensionalist stand.

REFERENCES

- [1] Deckert, M., “Quine, Strawson and logical truth,” *Philosophical Studies*, vol. XXIV (1973), pp. 52-56.

- [2] Harman, G., "Quine on meaning and existence, I," *Review of Metaphysics*, vol. XXI (1967), pp. 124-151.
- [3] Quine, W. V. O., *Philosophy of Logic*, Prentice-Hall, Inc., Englewood Cliffs, New Jersey (1970).
- [4] Quine, W. V. O., "Reply to Strawson," in D. Davidson and J. Hintikka (eds.), *Words and Objections*, D. Reidel Publishing Co., Dordrecht, Holland (1969).
- [5] Quine, W. V. O., *Word and Object*, The Massachusetts Institute of Technology Press, Cambridge, Mass. (1960).
- [6] Strawson, P. F., "Propositions, concepts and logical truths," *Philosophical Quarterly*, vol. VII (1957), pp. 15-25.

The Ohio State University
Columbus, Ohio