

Pietroski on possible worlds semantics for belief sentences

JOE LAU

On the possible worlds theory, propositions are identified with sets of possible worlds. A proposition p is true at a possible world w if and only if w is a member of p . How this theory can be integrated into a semantics for belief sentences is a notoriously difficult matter. Paul Pietroski has recently offered a proposal ('Possible worlds, syntax, and opacity', *Analysis* 53 (1993) pp. 270–80). The aim of this paper is to raise a worry and an objection. Pietroski accepts *the binary analysis* of belief sentences, as follows:

- (1) 'X believes that p ' is true exactly when X believes the proposition denoted by 'that p '.

He also assumes *omnidoxasticity*:

- (2) If X believes propositions $p_1 \dots p_n$, and $\{p_1 \dots p_n\}$ entails q , then X also believes q . (So in particular if X believes p , X believes every set of worlds that has p as a subset.)

I shall not challenge either assumption for the purpose of discussion. Given (1) and (2), the *problem of equivalence* arises if we take 'that p ' to denote the intension of ' p ', i.e. $\{w:p \text{ in } w\}$. So for example, since necessarily equivalent sentences have the same intension, the following two sentences are predicted to have the same truth-conditions, contrary to our semantic judgment:

- (3) Hilbert believes that $2 = 1$.
 (4) Hilbert believes that arithmetic is complete.

To avoid such and similar counterintuitive consequences, Pietroski proposes that,

- (5) 'that p ' denotes the set of worlds where (a) it is true that p , and (b) there is some true sentence s which is similar to ' p '. (Similar with respect to some relevant feature specified by the context, e.g. they have to be translations of each other.)¹

The central idea of this proposal is that *normally* the truth of a belief sentence requires the subject to have an appropriate metalinguistic belief. In contexts that are not 'normal', such as when we attribute beliefs to non-speakers, e.g. dogs, this requirement is dropped by omitting (5b). In other abnormal contexts, such as when mathematical beliefs are attributed, we drop (5a) instead. So consistent with our semantic intuitions, Pietroski's proposal does predict that (3) and (4) have different truth-conditions, since their that-clauses denote distinct propositions (6) and (7) respectively:

- (6) $\{w:\text{something similar to '2 = 1' is true in } w\}$
 (7) $\{w:\text{something similar to 'arithmetic is complete' is true in } w\}$

One major worry with this proposal is that what makes a context normal or abnormal is left rather unclear. We do know that in an abnormal context, either clause (5a) or (5b) is dropped. Still, this leaves open the question of *when* to adopt the normal or abnormal reading. As mentioned, Pietroski does provide some *examples* of abnormal contexts. So attributions of mathematical beliefs are instances of abnormal contexts where (5a) is dropped, and attributions to non-speakers are cases where (5b) is dropped. But this falls short of giving a general account of what makes a

¹ Pietroski provides a formal recursive definition of the denotation of a that-clause, but since his proposal does not deal with demonstratives, the 'disquotational' version is what his proposal comes down to.

context abnormal. Consider for example the belief report 'The confused undergraduate student believes that there are no thinking beings'. In this context we are neither attributing mathematical beliefs, nor attributing beliefs to non-speakers. Are we to say then, that the sentence should be given a normal reading, and thus both (5a) and (5b) apply? In that case the that-clause denotes the conjunction of $\{w:\text{there are no thinking beings in } w\}$ and $\{w:\text{there is a true sentence in } w \text{ that is similar to 'there are no thinking beings'}\}$. However, worlds where no thinking beings exist are presumably worlds where no languages are spoken, and for Pietroski this implies that there are no true sentences in these worlds. So on a normal reading the that-clause denotes the empty set. This predicts that the sentence would have to be false, since nobody believes the empty set. However, intuitively, there is no reason why the sentence cannot be true. To avoid this problem, Pietroski would have to say that this is also a case of abnormal context, but why should it? Sometimes Pietroski speaks of 'ignoring' (5a), or 'taking into account' (5b). Does it mean whether a context is normal or not is a matter of the speaker's intention? But this cannot be right. For surely I cannot make (3) and (4) come out to have the same truth-conditions, simply by intending that the context is normal. But if normality (or the lack of it) is not determined by the speaker's intention, then what else is relevant? Is it a matter of tacit conventions? But what might be the content of those conventions? We still need to fill in the blank in 'it is a matter of convention that a context with feature ___ is a normal context'. What Pietroski has yet to provide are motivated principles that distinguish normal from abnormal contexts.

Such a problem would be less pressing if belief sentences can always be assigned the correct truth-conditions under either the normal or abnormal reading. Unfortunately this is not the case. Consider again our monolingual German mathematician Hilbert, for which (4) is true. But now suppose Hilbert is told that 'arithmetic is incomplete' is a true English sentence, but he has no idea what it means. Hilbert still believes that arithmetic is complete, but taking his informer to be trustworthy, he comes to acquire the new and true belief that 'arithmetic is incomplete' is true. So he now believes the set K , $\{w:\text{'arithmetic is incomplete' is true in } w\}$.² Consider the sentence 'Hilbert believes that arithmetic is incomplete', which intuitively is false. However, Pietroski's theory predicts the opposite result. First, if we suppose that this is an abnormal context where (5b) does not apply, then the that-clause denotes the necessarily true proposition, of

² As pointed out by a referee, I assume that given X believes that s is true, it follows that X believes $\{w:s \text{ is true in } w\}$. This assumption is made by Pietroski himself (see *his* p. 275). Without such an assumption, there would also be no motivation in requiring that (5b) be dropped when attributing beliefs to non-speakers.

which K is a subset. Since Hilbert believes K , by (2) he also believes the necessarily true proposition, and so the sentence is true on this reading. On the other hand, if we adopt the other abnormal reading which drops (5a), or the normal reading, then the that-clause in both cases would denote the same metalinguistic proposition L , which is $\{w:\text{something similar to 'arithmetic is incomplete' is true in } w\}$. Regardless of what the relevant similarity relation is in the context, surely it has to be reflexive, and so K is a subset of L . As Hilbert believes K by (2) again he also believes L . The conclusion is that on all admissible readings, the theory predicts wrongly that 'Hilbert believes that arithmetic is incomplete' is true.

It will not do to respond by stipulating that similarity has to be non-reflexive. For presumably there are other non-German (e.g. French) sentences which are distinct from but are similar to 'arithmetic is incomplete'. If Hilbert believes any of them to be true the problem still arises. It will also not do to respond by rejecting (2), since it plays an essential role in Pietroski's proposal in explaining belief attributions to non-English speakers. Note also that more complicated counterexamples involving attributions of contingent beliefs can be constructed as well. Thus consider Lusina, a monolingual English speaker who believes that Twain wrote novels but that Clemens did not. Suppose she comes to believe of some sentence in French that it is true, but which unbeknownst to her translates into English as 'Clemens wrote novels'. (Perhaps she mistakenly thinks it means Clemens wrote *no* novels.) So she believes both $\{w:\text{Clemens/Twain wrote novels in } w\}$ and $\{w:\text{something similar to 'Clemens wrote novels' is true in } w\}$. By (2) Lusina also believes the conjunction of these two propositions, which on a normal reading, is the denotation of the that-clause in 'Lusina believes that Clemens wrote novels'. So this sentence is predicted to be true even though intuitively it still is false. I conclude that Pietroski's proposal fails to provide a plausible possible worlds semantics for belief sentences.³

The University of Hong Kong
Hong Kong
 jyflau@hkuxa.hku.hk

³ Thanks to Bob Stalnaker for comments on an earlier version of this note.