Is Yablo's paradox Liar-like?

JAMES HARDY

Stephen Yablo claims to have found a Liar-like paradox that does not depend on circularity or self-reference.[1] While I do not wish to dispute that what Yablo presents is indeed a paradox, I do wish to point out three important differences between Yablo's paradox and what is typically called the Liar Paradox. Here is the paradox as Yablo states it.

> Imagine an infinite sequence of sentences S_1, S_2, S_3, . . ., each to the effect that every *subsequent* sentence is untrue:
>
> (S_1) for all $k>1$, S_k is untrue,
> (S_2) for all $k>2$, S_k is untrue,
> (S_3) for all $k>3$, S_k is untrue, ...
>
> Suppose for contradiction that some S_n is true. Given what S_n says, for all $k>n$, S_k is untrue. Therefore (a) S_{n+1} is untrue, and (b) for all $k>n+1$, S_k is untrue. By (b), what S_{n+1} says is in fact the case, whence contrary to (a) S_{n+1} is true! So *every* sentence S_n in the sequence is untrue. But then the sentences *subsequent* to any given S_n are all untrue, whence S_n is true after all!

The most obvious difference between Yablo's paradox and traditional versions of the Liar is that Yablo's paradox is generated by considering an infinite collection of sentences (S_0, S_1, and for each n, S_n). If we restrict ourselves to a finite collection of the S_n, then no paradox arises. The upshot of this is that there is no first-order derivation of a contradiction from Yablo's premisses and the Tarski biconditionals.

The second difference is that Yablo's paradox requires an infinite number of instances of the T-schema. With traditional versions of the Liar we only need to appeal to a finite number of instances of the T-schema. Yablo's paradox will not go through if we limit ourselves to only a finite number of such instances. Suppose otherwise, then there is a proof of a contradiction from the Yablo sentences and some finite subset of the T-biconditionals for the Yablo sentences. Call this proof P.[2] Since P mentions only a finite number of the T-biconditionals, there is a greatest number m such that P mentions the T-biconditional for S_m. Now consider a model in which for all $n>m$, the T-biconditional for S_n does not hold. (Note that in

[1] Stephen Yablo, 'Paradox without Self-Reference', *Analysis* 53 (1993) 251–52.

[2] I'm using the term 'proof' here to include infinitely long proofs. Of course no finite proof from Yablo's premises could ever yield a contradiction.

ANALYSIS 55.3, July 1995, pp. 197–198. © James Hardy

this model 'T' will not mean true.) Further, for all $n>m$ let S_n not be in the extension of T. Thus $\neg T(S_n)$ holds for each $n>m$. But this is just what S_m says, so S_m holds in the model. Since we have the T-biconditional for S_m, $T(S_m)$ also holds in the model. Since $T(S_m)$ holds, then, for all $k<m$, S_k fails to hold. Since we have all of the relevant T-biconditionals we have $\neg S_k$ and $\neg T(S_k)$ for each $k<m$. So, in the model, the Yablo sentences before m do not hold while those after and including m do hold. The T-biconditionals hold for the Yablo sentences before and including m, but do not hold for the Yablo sentences after m. That this is a consistent model, shows that no contradiction follows from the Yablo sentences and any finite number of their T-biconditionals.

Finally, it turns out that even with all of the relevant T-biconditionals, the Yablo sentences are only omega inconsistent, not negation inconsistent. If there were a negation inconsistency, then, by compactness, that inconsistency would be confinable to a finite set of sentences. But no finite subset of the Yablo sentences are inconsistent.[3] What we do get in Yablo's paradox is an omega inconsistency. Suppose that each S_n is false. The falsity of S_1 implies that there is some n such that S_n is true, but it does not require that any particular S_n be true. There is no inconsistency between the falsity of S_1 and the falsity of any other finite set of S_n. The inconsistency on which Yablo's paradox feeds is between 'for some $n>1$, S_n is true', which is implied by the supposed falsity of S_1, and 'for all $n>1$, S_n is untrue' which seems to be a consequence of the fact that each S_n is false. But 'for all $n>1$, S_n is untrue' does not follow from the falsity of the various S_n without the further assumption of omega completeness. So Yablo's paradox is based on an omega inconsistency, whereas the traditional Liar is based on a negation inconsistency.

While Yablo's paradox does have some of the same sort of 'feel' to it that the Liar does, a closer examination reveals that the inconsistency underlying it is in some respects different than that which underlies the Liar. Is Yablo's paradox Liar-like? In some ways yes, and in other ways no.[4]

Indiana University
Bloomington, IN 47405-2601, USA
jhhardy@phil.indiana.edu

[3] However, an explicit contradiction could be deduced if we had, not the T-biconditionals, but the following generalization of them: $\forall n[S_n \text{ is true} \rightarrow \forall j>n(S_j \text{ is untrue})]$. The point is that while the Liar paradox may be seen as stemming from an inconsistency in the T-biconditionals themselves, Yablo's paradox may not.

[4] I would like to thank Anil Gupta for his helpful comments on both the substance and the execution of this paper.