Is Yablo's paradox non-circular?

Jc Beall

1. *Introduction*

Is there a non-circular liar paradox? If so, what is it? Stephen Yablo (1985, 1993) claims to have specified a particular example, *Yablo's paradox* as it is now called. Yablo's example, however, has been challenged. In particular, Graham Priest (1997) has argued that Yablo's paradox, though genuinely paradoxical, is not non-circular; and Priest's challenge applies across the board to any other allegedly non-circular Yabloesque paradox.[1]

As is often the case in philosophy, challenges breed challenges. The latest challenge comes from Roy Sorensen (1998), who has defended Yablo's claim of non-circularity in the face of Priest's objection. Evidently, Sorensen's defence has been thought by many to be successful; there has been little discussion, and no challenges, thereafter. In this paper I break the silence. My aim is to show that Sorensen's defence fails to address Priest's basic point; none of Sorensen's replies provides any reason for thinking that, contrary to Priest's challenge, Yablo's paradox is non-circular.

To make matters clear I shall present a new version of Priest's argument, a version which, I hope, expresses Priest's basic point in a clearer way than

[1] In the wake of Yablo's paradox have sprung a variety of other allegedly non-circular paradoxes – other Yabloesque paradoxes, as they are called, each being a variation on Yablo's original theme. See, for example, Goldstein's Yabloesque Russell paradox (1994), Beall's Yabloesque Curry (1999), and other examples cited by Sorensen (1998).

Analysis 61.3, July 2001, pp. 176–87. © Jc Beall

Priest himself presented it. With the new version in hand the failure of Sorensen's replies may be seen easily.

§2 briefly rehearses Yablo's paradox and presents (my version of) Priest's argument. §3, in turn, presents Sorensen's three basic replies and shows each to be wanting. §4 considers what are likely to be the two main objections to my discussion. §5 offers a very brief summary.

2. *Yablo's paradox and the case for its circularity*

In this section I briefly review Yablo's paradox and the charge of circularity advanced by Priest. I assume familiarity with the debate, and so the review is brief. (A review is also given in Sorensen 1998 and, with more formal details, Priest 1997.) The chief aim of this section is to present Priest's result in a new, simple, and (I hope) clearer light.

2.1. *Yablo's paradox and Priest's point*

A standard liar paradox involves a sentence that apparently says of itself only that it is not true.[2] One natural reaction to the liar paradox is to blame self-reference or, in the case of so-called 'loop liars', circularity. Another, more popular, response is to invoke some sort of hierarchy – a hierarchy of 'truth' predicates, as in Tarski 1944, a hierarchy of contexts, as in Burge 1979, or so on. None of these responses gets off the ground if there is a non-circular liar paradox. Is there such a thing? Stephen Yablo (1985, 1993) claims that there is.

The basic idea, as Sorensen (1998: 137) puts it, 'substitutes the cramped circularity of self-reference with the luxuriant linearity of an infinite series'. Specifically, we are to imagine a denumerable sequence of sentences each of which says only that the subsequent sentences are not true:

(s_0) For all k greater than 0, s_k is not true,
(s_1) For all k greater than 1, s_k is not true,
(s_2) For all k greater than 2, s_k is not true,
:
:

It would seem that we have hereby imagined not only a paradoxical sequence but also, more importantly, a *non-circular* paradoxical sequence. Given familiar T-schema and quantifier moves, the apparent sequence seems to be paradoxical. Moreover, given that apparently no sentence in our imagined sequence refers to itself, we also have a *non-circular* sequence. Such is Yablo's paradox.

[2] This version is often called *the strengthened liar*, a version, as Bas van Fraassen (1968: 147) put it, 'designed especially for those enlightened philosophers who are not taken in by bivalence'.

Appearances, however, are notoriously deceptive. There are a few misleading appearances at work in Yablo's original description, the most important of which concerns the appearance of non-circularity – that the imagined sequence, as Yablo (1993) put it, is 'not in *any* way circular'. In fact, there is circularity involved. For example, Yablo's original description appears to involve the T-schema, rather than a generalized satisfaction predicate (ranging over open sentences). As Priest (1997: 237–38) shows, however, the paradox, if well-defined at all, actually involves a predicate, 'no number greater than x satisfies this predicate', where the circularity is evident.[3]

The same point (in effect) can be seen another way. Sorensen (1998: 144–45) sums up the situation well:[4]

> As a finite thinker, Yablo can only generate his infinite sequence with a quantified expression of the form
>
> (s_n) For all k greater than n, s_k is not true.
>
> The need for this proposition is disguised by casual presentations that merely list the first few members and then recourse to a vague 'etc.', 'and so on', or '…'. This explicit s_n formulation is self-referential in the sense that s_n uses its own location in the sequence as a reference point to specify which statements are not true, i.e. the statements after s_n. Priest stokes the suspicion that this is a relevant sense of 'self-reference' by casting the point as a fixed point theorem.

And indeed, Priest does so cast his point.[5] Unfortunately, Sorensen does not fully appreciate Priest's point, as I shall argue below.

What, though, *is* Priest's point? On the surface Priest's point is merely that, *pace* Yablo, there is circularity involved in the (descriptive) specification of Yablo's paradox. Moreover, as Priest shows, specifying the paradox by recourse to infinitary languages, as per Forster (typescript), does not avoid the problem: circularity still emerges. Most generally, Priest's point, on the surface, is simply that any human description of Yablo's paradox – or, as Sorensen puts it, any such description given by a finite thinker – will involve circularity.

[3] For the formal details consult Priest's paper.

[4] Sorensen uses 'Y_k' where I've written 's_n'. Nothing hangs on this.

[5] Priest's point, of course, applies equally well to Sorensen's own version of Yablo's paradox, namely the Sorensenian queue paradox, and likewise to any Yabloesque construction. In Sorensen's queue paradox we are to imagine a denumerable queue of students, each thinking precisely the same thought: *Some of the students behind me are now thinking an untruth*. As Priest (1997: 240) points out, the circularity involved in the queue is evident: The relevant thought, t, is simply $t = \lambda x \forall y$(if y is behind x then $t(y)$ is not true), where $t(x)$ is x's thought.

How important is Priest's point? Perhaps so put the point appears to be less than important. Indeed, Sorensen himself concedes Priest's point so put, namely that any human description of Yablo's paradox will involve circularity. According to Sorensen, however, this is not 'a relevant sense of "self-reference"' (145). Sorensen's view is that while our specification of Yablo's paradox may be (indeed, is) circular, this does not provide reason for thinking that Yablo's paradox – the sequence so specified – is circular. On this matter, however, Sorensen is incorrect. Priest's point, I suggest, is much more significant than Sorensen lets on; the trouble, perhaps, is that it needs to be slightly recast.

2.2. *The case for circularity, recast*

I shall try to present Priest's point in a slightly different way, a way that seems to make it easier to see the failings of Sorensen's replies, which are discussed below (§3). In a nutshell, the argument may be put roughly as follows.

Let t be some term, and let $\delta(t)$ be the referent (or denotation) of t. We basically have two ways of fixing the reference of t: Demonstration or (attributive) Description. In the first case we need to see $\delta(t)$ or otherwise stand in close (spatial) proximity to $\delta(t)$, at which point we 'baptize' $\delta(t)$ and thereby fix the reference of t. This is the general situation involved in reference-fixing via demonstration. (The details, of course, are notoriously tricky; but for present purposes the general picture is sufficient.) The other way of fixing the reference of t, namely via (attributive) description, does not (apparently) require that we see $\delta(t)$. Instead, we simply let t denote whatever satisfies some description, D. (I here ignore cases where D fails to pick out a unique entity. These are philosophically important cases, but they are not relevant to the current debate at hand, given that all sides agree that 'Yablo's paradox' is determinately, uniquely referring.)

The foregoing division of reference-fixing is familiar enough. What is important for present purposes is the question: Have we fixed the reference of 'Yablo's paradox' by the first method (demonstration) or the second (description, attributive)?

Everyone, I think, will agree: we have not fixed the reference of 'Yablo's paradox' via demonstration. Nobody, I should think, has seen a denumerable paradoxical sequence of sentences, at least in the sense of 'see' involved in uncontroversial cases of demonstration.[6] Perhaps one might protest that we have seen Yablo's paradox; we have seen tokens of some of its constituent parts, as displayed, for example, in §2.1; and seeing just this

[6] In §4 I briefly consider the line according to which we 'see' the sequence via some sort of mental 'sight' or the like. For now, however, I concentrate only on standard, uncontroversial cases of reference-fixing demonstrations.

much, the suggestion goes, is sufficient for the required 'baptism', which thereby affords fixing the reference of 'Yablo's paradox'.[7] This suggestion, while interesting, cannot be maintained easily, at least without much more explanation. To begin, there are infinitely many distinct sequences whose first few tokens appear exactly as displayed in §2.1. Which one is being baptized? Answering this question requires further specification; and further 'seeing', presumably, will not do the work – work that requires specifying a unique referent of 'Yablo's paradox'. Moreover, as Sorensen himself concedes, the 'presentation' displayed in §2.1 is merely casual, depending on a vague 'etc.', 'and so on', or '...'. 'Seeing' just this, then, would seemingly be seeing not enough – or not seeing clearly enough, as it were. What is required, if demonstration is to do the trick, is that we see the sequence (or a token of it) in full; otherwise, we run the risk of not fixing the reference of 'Yablo's paradox' on a unique sequence (if a well-defined infinite sequence at all). The going suggestion, then, does not seem to be plausible. If the suggestion is to find plausibility, a lot more needs to be said by defenders of the line, a lot more than has been said by defenders of non-circularity thus far, including Sorensen and Yablo. For now, however, I assume, as said, that most philosophers (including, I think, both Sorensen and Yablo) will agree that the reference of 'Yablo's paradox' has not been fixed by demonstration.

If the foregoing is correct, then the issue becomes clear: if we have fixed the reference of 'Yablo's paradox' at all, then we have fixed the reference of 'Yablo's paradox' via (attributive) description. But, now, the upshot of Priest's point is plain: Priest has shown that any description *we* employ to pick out (or otherwise define) a Yabloesque sequence is circular; this much Sorensen concedes. From here, however, it is a small step to the circularity of the sequence itself. We are fixing the reference of 'Yablo's paradox' via (attributive) description, which means that 'Yablo's paradox' denotes whatever satisfies the given reference-fixing description. The situation, however, is this: that the satisfaction conditions of our available reference-fixing descriptions require a circular satisfier – a sequence that involves circularity, self-reference, a fixed point. Given all this, it follows that the reference of 'Yablo's paradox' is circular. The upshot of this is apparently missed by Sorensen; the upshot is that, unless we find some other way of fixing the reference of 'Yablo's paradox', we are stuck fixing it on a circular sequence – a sequence containing fixed points, self-reference, etc.

Priest's point, then, is significant indeed. Cast in the light of how we fixed the reference of 'Yablo's paradox' the point places a heavy burden on defenders of a non-circular Yablo paradox. All parties agree that 'Yablo's

[7] Mark Colyvan suggested as much in conversation, though he is not committed to the suggestion.

paradox' is a well-defined term, picking out a unique sequence. As such, the reference of the term had to be somehow fixed. Demonstration seems to afford little help, as suggested above, and is at any rate an implausible candidate, at least without further explanation. This leaves the second reference-fixing technique: (attributive) description. And that is the rub. Given that, as Priest has shown, any description of Yablo's paradox – and a fortiori any reference-fixing description – requires, for unique satisfaction, circularity, we have it that we are talking about a circular sequence when we use 'Yablo's paradox' (or the like), on pain of talking about no sequence at all.

Accordingly, the burden is on Sorensen, Yablo, or other defenders of non-circularity to show that 'Yablo's paradox' does not denote a circular sequence. This, I suggest, is the basic issue in the light of which Sorensen's replies should be read. So read, however, none of Sorensen's replies succeed in overturning Priest's (recast) point. That this is so is argued below.

3. *Sorensen's replies: God, demonstrations, and definite descriptions*

Sorensen (1998) attempts to answer Priest's point and show that what we are talking about when we are talking about Yablo's paradox is a non-circular (infinite, paradoxical) sequence. He gives three closely related replies. In each case, however, his reply simply sidesteps the basic issue. In what follows I treat each of Sorensen's replies in turn, showing in each case that they fail to defend against Priest's basic point (as presented in §2.2). In each case the problem is virtually the same: Sorensen's remarks, though true, are seemingly beside the point. To make matters simple I shall quote Sorensen's relevant remarks at length, and then simply point out the respective failures to address the issue at hand (again, as presented in §2.2.).[8]

3.1. *God*

Sorensen's first defence invokes God or the abilities of infinite beings. As Sorensen puts it:

> Priest's objection underestimates the resources available to those who wish to specify sequences. An infinite being could enumerate a denumerable list of sentences. For example, God could write the first sentence during the first minute, the second sentence after following 30 seconds [*sic*], the third in the following 15 seconds, and so on. By writing faster and faster, God could finish the sequence in two minutes. Since we finite beings know that the Yablo sequence is paradoxical for

[8] The chief objection to my following discussion will be a charge of unfair play, that in effect I've changed the target and merely shown that Sorensen has missed the new target, the one to which he wasn't aiming. I address this concern in §4.

> the infinite being, we know that the Yablo sequence is paradoxical *simpliciter*. Our use of a self-referential specification is merely a useful heuristic. (145)

The trouble with gods: Sorensen is correct about the resources available to an infinite being; however, nothing he says addresses the issue at hand. What is needed is a reason for overturning Priest's case, presented in §2.2, that the referent of 'the Yablo sequence' or 'Yablo's paradox' is non-circular. Granted, a god can enumerate an infinite sequence in the given way; that is not at issue. The issue is: What sequence are we talking about when we ask the god to enumerate it? Since, as in §2.2, we have fixed the reference of 'Yablo's paradox' by a description satisfied only by a circular sequence, we will be asking the god to enumerate a circular sequence. As Sorensen says, we do indeed know that the given Yablo sequence is paradoxical for the infinite being; that is not at issue. The basic issue concerns which sequence is paradoxical. Given the way we have fixed the reference of 'Yablo's paradox', the infinite being, just like us, is considering a circular paradox.

Sorensen might reply that if we had the god's abilities we could fix the reference of 'Yablo's paradox' (or any other term, for that matter) on a denumerable paradox free of circularity. To reply as such, however, seems to ignore the problem. As in §2.2, the upshot of Priest's point is that, for all we have been told, there simply is no (well-defined) denumerable paradox free of circularity, and a fortiori no such paradox to which even a god could refer. Perhaps gaining a god's abilities would illuminate such a paradox; but, for all that Sorensen has said, that remains an open question. For now, we are stuck denoting that sequence on which *our* terms have been fixed – Yablo's paradox, which is circular.

3.2. Demonstrations

I have argued that, given our reference-fixing situation, we are talking about a circular sequence when we talk about Yablo's paradox. What I am assuming, here, is that we have fixed the reference of 'the Yablo sequence' or 'Yablo's paradox' via the given description (to which Priest's theorem applies), as discussed in §2.2. Sorensen's second argument enters on this point. Sorensen's second defence invokes the common idea of demonstrations, discussed in §2.2, where these serve to specify some entity without recourse to descriptions. Sorensen writes:

> Sequences can be specified demonstratively instead of descriptively. A methodologically scrupulous experimenter selects his random number sequence by randomly flipping to a page of the Rand Corporation's classic *A Million Random Digits*. If I see a queue of students, I can specify a sequence just by pointing at the queue – even if the queue is

> infinitely long. God need not have set up the queue. No one need be able to construct a roster. If a finite being wishes to describe which sequence is my demonstratum, then his description must be recursive and so self-referential in a sense. But the finite being's self-referential description of the demonstratum fails to make any component of the original demonstration (demonstrator, act of demonstration, demonstratum) self-referential. (145)

The trouble with demonstrations: Once again, Sorensen is correct about demonstrations, which, as he says, may be used to fix the reference of our terms. As he says, some sequence, σ, may not be circular in any way, even though any description of σ may involve circularity. The demonstration, however, allows us to denote the sequence without circularity. All of this is perfectly correct.[9] The question, put bluntly, is: So what? Priest's point, at least as given in §2.2, in no way conflicts with Sorensen's claims about demonstrations. Sorensen's claims seem to be entirely beside the point. He is right: if I see a queue of students (even an infinite one, etc.) I can specify the sequence just by pointing. This is true but irrelevant. If I see an infinite, non-circular paradoxical sequence, then I can point to it and thereby fix the reference of 'Yablo's paradox'. But neither I nor anyone else has seen such a sequence; and so Sorensen's second point seems to be beside the point.[10]

Once again, Sorensen might reply by saying that if, contrary to fact, we *had* seen a denumerable paradoxical sequence then we would be able to avoid the problem posed by Priest. This is correct, but it does not begin to make the required case: namely, that we have seen such a sequence or, more to the point, that we have fixed the reference of 'Yablo's paradox' in such a way that avoids Priest's point.

3.3. Descriptions

Sorensen's final, and perhaps central, reply invokes the work of Donnellan (1966) and Kaplan (1989), which, as Sorensen (1998: 148) says, 'lets us distinguish between attributive and referential uses of self-referential descriptions'. Sorensen notes that self-referential (or otherwise circular) descriptions can denote non-circular expressions, just as 'Any expression with five words', which is circular, denotes 'The cat on the mat', which is not circular. Sorensen writes as follows:

[9] Well, there is some question about whether demonstrations really rid us entirely of the need for descriptions, but I set this aside, just as Sorensen's reply does. For purposes of argument I grant the strongest case, namely that demonstrations rid us entirely of the need for descriptions.

[10] Again, in §4 I briefly address the invocation of mental 'sight', as it were, which seems to be the last available option for defenders of non-circularity.

> If I set up a sequence with a self-referential description, it does not follow that the *content* of what I specified is self-referential. 'The cat is on the mat' is specified with the descriptive phrase 'any sentence that has fewer words than this very sentence' but it is not thereby self-referential. Likewise, I may specify a sequence with a self-referential sentence without the sequence being thereby self-referential. This holds even if *my* only means of specifying the sequence is self-referential. (1998: 148)

Sorensen's main point, as he puts it, is simply this: 'Reference fixing sentences do not pass their properties (contingency, vagueness, indexicality, etc.) on to the sentences they rigidly designate. The architecture of a description does not mould the structure of what it describes – even if that description is describing other descriptions' (1998: 148).

The trouble with Donnellan and Kaplan: Once again, much of this is true but irrelevant. Priest's point in no way requires that the properties of reference-fixing descriptions pass on their own properties to the given referent. Priest's point, rather, is that in the Yablo case our only handle on the referent of 'the Yablo sequence' or 'Yablo's paradox' is via a description that is satisfied by a circular sequence. The point is not argued from the strong (and implausible) premiss that for any reference-fixing description, D, any property of D is a property of x, where x (uniquely) satisfies D. The point, rather, is that any description, D, used to fix the reference of 'Yablo's paradox' is such that D's satisfaction conditions require that the (unique) satisfier is circular (contains self-reference, a fixed point, etc.). The trouble with Sorensen's going reply is that, as with the others, it misses Priest's basic point.

3.4. *The problem, in summary*

I have argued that none of Sorensen's replies addresses the central issue at hand. As above, the problem with Sorensen's replies is not so much that they are incorrect; rather, they are simply off the mark. What needs to be shown is that, contrary to Priest's point (as put in §2.2), we have reason to believe that when we use 'Yablo's paradox' (or the like) we are talking about a non-circular paradox. Nothing Sorensen has said begins to provide such reason, and that is the main problem with his defence against Priest's point.

4. *Objections and replies*

There are two chief objections that are likely to be raised against this discussion. The first concerns fairness to Sorensen; the second concerns a neglected option that might be invoked in defence of non-circularity. I consider each in turn.

4.1. *Sorensen's target*

One objection to the foregoing discussion is a charge of unfair play. I have argued that Sorensen's replies to Priest leave Priest's point entirely intact. The problem, as said, is that Sorensen's replies, though on the whole true, simply miss the main target. The current objection is that I have cheated by changing the target; Sorensen's replies were aimed at another target, not the one to which I have taken them to be aimed.

Reply: I think that there is truth in the charge. Sorensen's defence seems to be aimed only at Priest's superficial point, the point that any human *description* of Yablo's paradox is essentially circular. According to Sorensen, circularity involved only *in description*, as opposed to content or the described, is not relevant to Yablo's paradox; indeed, it is 'off the record' (1998: 148). Taking only this much as a target, perhaps Sorensen's replies hit the mark. What I hope is clear, however, is that there is a more fundamental target that needs to be hit; and, as §3 shows, none of Sorensen's replies begin to hit the more fundamental target.

Perhaps, then, I am guilty of changing targets behind Sorensen's back, and guilty, in turn, of criticizing his missing the new target – a target of which he was not aware. What is important for the overall debate is that the new target be displayed, a target that, for now, has yet to be touched by defenders of non-circularity. The new target is that of §2.2, Priest's point recast.

4.2. *Demonstration and mental eyes*

There is another objection which is likely to arise. The objection is aimed at the claim that we have not fixed the reference of 'Yablo's paradox' via demonstration. One option available to Sorensen, Yablo, and anyone else who might want to argue that Yablo's paradox is non-circular, is as follows. They might say that in fact we have fixed the reference of 'Yablo's paradox' by the first method (demonstration). The idea is that while we have not seen the sequence in the normal, causal sense of 'see', we have nonetheless seen the sequence in a mental eye (as it were). Such seeing, as the story might go, is sufficient for fixing the reference of 'Yablo's paradox' on the non-circular paradox we saw; and the term is uniquely so fixed because we have seen the sequence in its entirety – or at least enough of it to rule out rival lookalikes.

Reply: I do not have knockdown arguments against this sort of line. I simply do not find the line to be plausible. In many ways, however, my own reservations about the given line are very much beside the point. I grant that if the reference of 'Yablo's paradox' can be fixed (uniquely) without recourse to description, then the case for circularity crumbles quickly. If mental eyes can do the trick, then so be it. At this stage, however, nobody has begun to invoke the mental eye defence, let alone explain it sufficiently. At the very least, much more needs to be said than what Sorensen actually says in his paper, if this 'mental eye' story is the primary case for believing in a non-circular liar paradox. Until then, we remain without good reason for thinking that 'Yablo's paradox' has been fixed by anything other than essentially circular descriptions, and so, for reasons given throughout, we remain without good reason for thinking that Yablo's paradox is anything other than circular.

5. *Summary*

My aim, here, was just to show that nothing Sorensen says provides us with reason to reject Priest's basic point. I have attempted to support this claim by showing that nothing Sorensen says begins to address the basic issue at hand. Unless Sorensen (or Yablo, or others) wishes to claim that the reference of 'Yablo's paradox' has been fixed by demonstration-via-the-mind's-eye, then we remain without reason for thinking that Yablo's paradox is a non-circular paradox. If the 'mind's eye' reply is to be given, then a lot more needs to be said. For now, Yablo's paradox is indeed a paradox, and an interesting one at that; but what *it* is remains circular.[11]

University of Connecticut
344 Mansfield Road, U-54
Storrs, CT 06269-2054, USA
beall@uconn.edu

References

Beall, Jc. 1999. Completing Sorensen's menu: a non-modal Yabloesque Curry. *Mind* 108: 737–39.

Burge, T. 1979. Semantic paradox. *Journal of Philosophy* 76: 169–98.

Donnellan, K. 1966. Reference and definite descriptions. *Philosophical Review* 75: 281–304.

Forster, T. Typescript. The significance of Yablo's paradox without self-reference. (Available from http://www.dpmms.cam.ac.uk.)

Goldstein, L. 1994. A Yabloesque paradox in set theory. *Analysis* 54: 223–27.

Kaplan, D. 1989. Demonstratives: an essay on the semantics, logic, metaphysics, and epistemology of demonstratives and other indexicals. In *Themes from Kaplan*, ed. J. Almog, J. Perry, and H. Wettstein, 481–614. New York: Oxford University Press.

[11] For comments on earlier drafts I thank Michael Clark, Mark Colyvan, Hartry Field, Jay Garfield, Laurence Goldstein, Graham Priest, Mike Resnik, Greg Restall, Roy Sorensen and Joel Stafford.

Priest, G. 1997. Yablo's paradox. *Analysis* 57: 236–42.
Sorensen, R. 1998. Yablo's paradox and kindred infinite liars. *Mind* 107: 137–55.
Tarski, A. 1944. The semantic concept of truth and the foundations of semantics. *Philosophy and Phenomenological Research* 4: 341–76.
van Fraassen, B. 1968. Presupposition, implication and self-reference. *Journal of Philosophy* 65: 136–52.
Yablo, S. 1985. Truth and reflection. *Journal of Philosophical Logic* 14: 297–349.
Yablo, S. 1993. Paradox without self-reference. *Analysis* 53: 251–52.

ANALYSIS 61.3, July 2001, pp. 187–93. © Anthony Brueckner