The ‘mental eye’ defense of an infinitized version of Yablo’s paradox
Selmer Bringsjord & Bram van Heuveln

1. Introduction

J.C. Beall (2001) argues that Yablo’s (1993) paradox, despite Sorenson’s (1998) defense to contrary, is circular. The crux of Beall’s case is the view that we have but two ways to fix the referent \( \delta(t) \) of some term \( t \): Demonstration or Description. In the former case we essentially need to “see” \( \delta(t) \); in the latter, we let \( t \) denote that which satisfies some description \( D \). Beall spends the bulk of his paper attempting to show that there is no non-circular description \( D \) of Yablo’s paradox, and we think he succeeds. But what about the Demonstration route? Beall, by his own admission, doesn’t completely close this route off. He says (2001: 179) that “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.” This comment leaves open the possibility that there is a determinate and defensible (but controversial) sense of ‘see,’ seeing\(_m\), on which Yablo’s paradox can be seen — and indeed in §4.1 of his paper, “Demonstration and mental eyes,” Beall does discuss this possibility:

The idea is that while we have not seen the sequence [at the heart of Yablo’s paradox] in the normal, causal sense of ‘see,’ we have nonetheless seen the sequence in a mental eye (as it were). (2001: 185)

Beall grants that he has no knockdown argument to rule out this possibility. He simply rests content to point out that “nobody has begun to invoke the mental eye defense, let alone explain it sufficiently.” (2001: 186) We provide this defense herein. The defense includes both a refutation of Priest’s (1997) argument that even infinitary versions of Yablo’s paradox are circular, and our attempt to exploit our seeing\(_m\) of an infinitized Yablo in order to approximate it (and the paradox of the gods, or “new Zeno,” introduced to Analysis readers by Priest (1999)) in finitary terms.

2. An infinitary version of Yablo’s paradox

The first step is to put on display that which is to be seen\(_m\), namely, an infinitary version of Yablo’s paradox, to which Priest (1997) has pointed, and erroneously declared to be circular:¹

Recall the familiar natural numbers \( \mathbb{N} = \{0, 1, 2, \ldots \} \). With each \( n \in \mathbb{N} \) associate a sentence as follows, using a truth predicate, \( T \):

\[
\begin{align*}
s(0) &= \forall k (k > 0 \rightarrow \neg T(s(k))) \\
s(1) &= \forall k (k > 1 \rightarrow \neg T(s(k)))
\end{align*}
\]

¹In the articles discussing Yablo’s paradox, writers refer to an infinitary version of Yablo’s paradox. For example, both Priest (1997) and Beall (2001) refer to such a formulation in an unpublished note by Forster, but the formulation simply isn’t there. Forster provides only the one-line kernel of an \( \mathcal{L}_{\omega_1\omega} \)-based definition of our function \( s \). Likewise, Priest (1997) says that Hardy (1995) provides an infinitary, \( \omega \)-rule-based version of Yablo’s paradox, but Hardy only proves, indirectly, that Yablo’s paradox entails \( \omega \)-inconsistency. In order to set the record straight and advance the debate, we specify an infinitary version of Yablo’s paradox, expressed in the “background” logic that allows for meta-proofs regarding infinitary logical systems like \( \mathcal{L}_{\omega_1\omega} \) (see note 7). This system is presented in encapsulated form in (Ebbinghaus, Flum, & Thomas 1984), from which the student interested in infinitary logic can move to (Karp 1964), then to (Keisler 1971), and then to (Dickmann 1975).
\[ s(2) = \forall k (k > 2 \rightarrow \neg T(s(k))) \]
\[ s(3) = \forall k (k > 3 \rightarrow \neg T(s(k))) \]

\[ \vdots \]

Expressed with help from the infinitary system \( L_{\omega_1 \omega} \), we can say that
\[ s(0) = \bigwedge_{k>0} \neg T(s(k)), s(1) = \bigwedge_{k>1} \neg T(s(k)), s(2) = \bigwedge_{k>2} \neg T(s(k)) \ldots \]

Next, suppose that \( T(s(0)) \). From this it follows immediately that \( \neg T(s(1)) \land \neg T(s(2)) \ldots \), which in turn implies by conjunction elimination in \( L_{\omega_1 \omega} \) that \( \neg T(s(1)) \). But in addition, if \( T(s(0)) \) is true, it follows again that \( \neg T(s(1)) \land \neg T(s(2)) \ldots \), and hence that \( \neg T(s(2)) \land \neg T(s(3)) \ldots \), which implies that \( T(s(1)) \). By reductio we can infer \( \neg T(s(0)) \).

The same indirect proof can be given to show \( \neg T(s(1)), \neg T(s(2)), \neg T(s(3)), \ldots \). Hence we can infer by the \( \omega \)-rule

\[ \begin{array}{c}
\alpha(1), \alpha(2), \ldots \\
\hline
\alpha(n) \\
\end{array} \]

that
\[ (\ast) \quad \bigwedge_{k \in \mathbb{N}} \neg T(s(k)). \]

Hence \( \neg T(s(1)), \neg T(s(2)), \neg T(s(3)), \ldots \), that is, \( T(s(0)) \). But \( \neg T(s(0)) \) follows from \( (\ast) \) — contradiction.

Note that it will not do to ignore the fact that this is an infinitary paradox expressed with indispensable help from \( L_{\omega_1 \omega} \). As such, it cannot be reduced to finitary, set-theoretic reasoning on which the paradox might be claimed to evaporate, or to be trivial. For example, someone might be tempted to view our version of the paradox (as well as, by implication, previous versions, including the original) as at bottom the attempt to define a set having the property of being such that:

\[ \ast \quad \forall n (n \in S \leftrightarrow \forall m > n (m \notin S)), \]

where a standard interpretation is presupposed. There is of course no such set (as a simple proof by cases on a partition of the power set of \( \mathbb{N} \) will reveal) — but that is irrelevant; this is so for two reasons. First, infinitary constructions in \( L_{\omega_1 \omega} \), as a matter of mathematical fact, cannot be reduced to constructions in FOL. The second reason our version of Yablo’s paradox cannot be reduced to simple set-theoretic assertions is simply that the paradox is not about sets. Rather, it is irreducibly about propositions referring to other propositions, truth, and so on. After all, even the simple Liar cannot be reduced to set-theoretic schemes analogous to \( (\ast) \).³

²Proofs of this are easy to come by. E.g., “finitude” can be captured effortlessly in \( L_{\omega_1 \omega} \) via
\[ \bigvee_{n < \omega} \exists x_1 \ldots \exists x_n \forall y (y = x_1 \lor \ldots \lor y = x_n); \]
i.e., an interpretation (in the standard model-theoretic sense) satisfies this formula if and only if the interpretation is finite. It is easy to prove that no such formula can be expressed in FOL.

³Consider this version of the Liar:
\[ p_n := \text{The sentence below is false.} \]
3. Why this infinitary version surmounts Priest

Of course, without the $\omega$-rule, this reasoning doesn’t fly. As Priest (1997) explains, if instead of an infinity of proofs, each of which establishes the relevant instantiation of $\alpha(1), \alpha(2), \ldots$, finitary reasoning expressible in first-order logic ($\mathcal{L}_I$) is given, a version of the $T$-schema that applies to formulas containing free variables must be invoked, and this will of necessity involve the notion of satisfaction. For example, such constructions as

$$S(n, \dot{s}) \rightarrow \forall k (k > n \rightarrow \neg Ts_k)$$

would be needed, where $S$ is the two-place satisfaction relation between numbers and the “predicate” $\forall k (k > n \rightarrow \neg Ts_k)$. As Priest shows, this leads to transparent self-referentiality when $\dot{s}$ is a predicate of the form $\forall k (k > n \rightarrow \neg S(k, \dot{s}))$ where $\dot{s} = \forall k (k > n \rightarrow \neg S(k, \dot{s}))$.

But the key question is: When the paradox is set out in the infinitary form we’ve given, is it still circular? Priest answers in the affirmative. He claims that “the circularity has nothing to do with the argument as such; it arises in the structure of the situation” (1997: 239). But without satisfaction to plague the version we’ve given, how exactly is it that circularity is at the heart of this formulation? Here is the crucial passage:

The function $s$ is defined by specifying each of its values, but each of these is defined with reference to $s$. (As a glance at [B&H’s] formulation suffices to demonstrate.) It is not the function $s$ that is a fixed point. $s$ is the function which, applied to any number, gives the claim that all claims obtained by applying $s$ itself to subsequent numbers are not true. Again, the circularity is patent. Note the role that the infinite regress is playing here. If the regress grounded out in some claim not concerning the sequence, then $s$ could be defined recursively, and it would not require a circular construction to define it. But the sequence is infinite; and it does. (Priest 1997: 239).

But this is really a stunning mistake. Priest simply overlooks the very possibility our formulation is designed to make explicit: an ungrounded infinite regress without circularity. This is exactly what we have. The general definition of our function $s$ is

$$s(n) = \bigwedge_{k > n} \neg T(s(k)),$$

and expanding, with variables kept distinct, yields

$$s(n) = \bigwedge_{k_1 > n} \neg T(\bigwedge_{k_2 > k_1} \neg T(\bigwedge_{k_3 > k_2} \neg T(s(\ldots k_3 \ldots)))).$$

This is an infinite construction, but it isn’t circular. The function $s$ is, to use Priest’s words when talking of Forster’s formulation, “getting in on both sides of the act,” but so what? We knew at the outset that in the definition of our function $s$ the symbol ‘$s$’ turns up on both sides, but that

$p_m :=$ The sentence above is true.

It would obviously be wrong to say, following (*), that the Liar evaporates or is trivial because there is no set $S$ of two natural numbers such that $n \in S \leftrightarrow m \notin S$ and $m \in S \leftrightarrow n \in S$. If this were not wrong, then nearly everything written about the Liar would be swept away as utterly superfluous. In particular, there would be no reason to create clever systems in which the likes of $p_n$ and $p_m$ can be represented without any contradiction. Likewise, again, if the set-theoretic reduction of infinitary Yablo to (*) was correct, Yablo’s paradox should have been cast aside before publication: all the words in Analysis and Mind on it would be almost downright silly. E.g., Priest’s (1997) ingenious objection to it would be wholly superfluous; ditto for Beall’s (2001).
doesn’t immediately imply circularity. It was the ordinary concept of satisfaction that gave rise to circularity in the first-order version of the paradox; that concept is absent in our formulation. While it’s true that standard inductive definitions “bottom out,” nothing says that if they don’t bottom out they must be circular. That’s something that must be demonstrated on a case-by-case basis. In the case of our formulation, no such demonstration is possible, because the definition of our function simply implies an infinite construction that follows an effortlessly seen pattern forever, without circularity. The point here is quite unexceptionable. Countless infinitely expanding patterns, for example the innocuous \((1, (1, (1, (1, \ldots)))\)), can be expressed as functions that aren’t circular.

4. The ‘mental eye’ defense

We begin by overthrowing Priest’s argument for the view that the \(\omega\)-rule can’t be used by humans. Here’s the argument adapted slightly so as to target our version of the paradox:

[The suggestion that we can use the \(\omega\)-rule] would be disingenuous, though. As a matter of fact, we did not employ the \(\omega\)-rule, and could not have. The reason we know that \(\neg Ts_n\) is provable for all \(n\) is that we have a uniform proof, i.e., a proof for variable \(n\). Moreover, no finite reasoner ever really applies the \(\omega\)-rule. The only way that they can know that there is a proof of each \(\alpha(i)\) is because they have a uniform method of constructing such proofs. And it is this finite information that grounds the conclusion \(\forall x \alpha(x)\). (Priest 1997: 239).

This is not a strong argument, to put it mildly. What premises and inferences therefrom support the proposition that we didn’t and can’t use the \(\omega\)-rule? Priest simply says that “as a matter of fact” this proposition is true; he thus gives us a bald petitio. In the proof we gave above, we did not reason over variables for natural numbers, as can be seen by inspection. However, if we’re charitable, we can view the second part of the quote as an argument, viz., an enthymematic version of

1. No finite reasoner ever really applies the \(\omega\)-rule.
2. We (or B&H) are finite reasoners.
3. \(\therefore\) We (or B&H) don’t ever really apply the \(\omega\)-rule.

This argument would appear to be formally valid; at any rate, we concede that it is. But are the premises true? Well, what is a finite reasoner, exactly? One answer, flowing from a computationalist conception of mind, is to say that a finite reasoner is at bottom a Turing machine or an equivalent, or a computing machine of less power (e.g., a linear bounded automaton, essentially a Turing machine with a finite tape). We can then appeal to well-known theorems to support (1): a Turing machine, rendered in declarative form, corresponds to no more than finite, standard deduction from finite axiom sets, all expressed and carried out in \(L_1\). So far so good. But why, then, are we supposed to affirm the second premise, (2)? After all, arguably the pivot around which philosophy of mind revolves, these days, is whether or not we are Turing machines (or “less”). Specifically relevant is the fact that there are many arguments, some book-length, for the falsity of the view that we are Turing machines (e.g., to pick close to home: Bringsjord 1992, Bringsjord & Zenzen 1997). The point isn’t that such arguments should be assumed here to be sound; not at all. The point, rather, is that in light of such arguments, Priest is in no position to simply assume (2) — and hence he hasn’t derailed the infinitary version of Yablo’s paradox.
Beall is no better off than Priest. As we have already noted, Beall admits that the mental eye defense at least has a fighting chance. And Beall’s own reasons for being skeptical about this defense are weak. Recall that Beall (2001: 179) writes: “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.” This is circumspect, but if we read it as a subtle nod toward an argument like “Nobody can see a denumerable sequence of sentences, because such a sequence is infinite, and we, as finite reasoners, can at most see finite sequences.” Beall is presupposing that we are finite reasoners, and so offers us nothing better than Priest’s unconvincing case.4

But what can be said in favor of the view that we can see, in some determinate sense (i.e., see_m), the elements in the infinitary reasoning we’ve presented? Let’s first address this question by putting it in line with Priest’s argument: What can be said in favor of the view that we are infinite reasoners? Well, if it turned out that we were not Turing machines, but hypercomputers instead, then we would by definition have the capacity to use the ω-rule. It’s important to note here that hypercomputers don’t in any way “cheat” when it comes to this rule. Some readers may be inclined to extend Priest’s talk of merely using versus really using the ω-rule (see note 4), and stubbornly declare that obviously we can’t make essential use of this rule, but can only see that the infinitely many premises were true by (and here our opponent reiterates Priest’s claim) some uniform method. This line against us is a dead end, for hypercomputers, unlike Turing machines and their equivalents (and lesser systems), can make essential use of the ω-rule; this is just a brute mathematical fact. If we are not Turing machines, but hypercomputers instead, it follows immediately that we can make essential use of the ω-rule.

Now, just as there are an infinite number of mathematical devices equivalent to Turing machines (Register machines, the λ-calculus, abaci, . . . ; these are all discussed in the context of an attempt to define computation in Bringsjord 1994), there are an infinite number of ways of specifying hypercomputers. Fortunately, a small proper subset of these specifications dominate the literature; in fact, three kinds of hypercomputational devices — analog chaotic neural nets, trial-and-error machines, and Zeus machines — are generally featured in the literature. Due to lack of space, we discuss only Zeus machines here.5

Zeus machines are named for the character Zeus, described by Boolos and Jeffrey (1989). (The machines were introduced and discussed long before this, apparently first by Bertrand Russell (1915, 1936). They have sometimes been known as ‘Weyl machines’ because of (Weyl 1949). Recently Copeland (1998) has called them ‘accelerated Turing machines.’) Zeus is a superhuman creature

4 We have actually given Priest a bit of a free ride: For notice that Priest says that finite reasoners don’t ever really apply the ω-rule. This implies that there is a sense according to which finite reasoners do apply this rule; and that sense is in fact easy to point to. The idea is doubtless that when a finite reasoner has some kind of finite description of an infinite sequence (e.g., α(1), α(2), . . . can itself be seen as such), and that description figures in concluding α(n), the finite reasoner can be said to have used the ω-rule . . . but not really, since really using the omega rule requires one to have a description of each of the individual entries in this series, and that’s something finite reasoners apparently don’t have. But why should we assume that it isn’t sufficient that finite reasoners use the ω-rule, but not really use it? More specifically, why should we assume that Priest is right that to use it simpliciter is to rely on “uniform methods”? 

5 A concise survey of hypercomputers can be found in the first chapter of (Siegelmann 1999), after which analog chaotic neural nets are characterized (for a more compressed characterization of such nets see Siegelmann and Sontag 1994). For cognoscenti, analog chaotic neural nets are (artificial) neural networks allowed to have irrational numbers for coefficients. For the uninitiated, analog chaotic neural nets are perhaps best explained by the “analog shift map,” presented in (Siegelmann 1995) and (Bringsjord 1998). Trial-and-error machines have their roots in a paper by Hilary Putnam (1965), and one by Mark Gold (1965); both appeared in the same rather famous volume and issue of the Journal of Symbolic Logic. Trial-and-error machines have the architecture of Turing machines (read/write heads, tapes, a fixed and finite number of internal states), but produce output “in the limit” rather than giving one particular output and then halting.
who can enumerate \( \mathbb{N} \) in a finite amount of time, in one second, in fact. He pulls this off by giving the first entry, 0, in \( \frac{1}{2} \) second, the second entry, 1, in \( \frac{1}{4} \) second, the third entry in \( \frac{1}{8} \) second, the fourth in \( \frac{1}{16} \) second, \ldots, so that, indeed, when a second is done he has completely enumerated the natural numbers. It is an easy matter to translate this behavior into mathematically precise adaptations of the standard account of Turing machines (Bringsjord 2001). Even in the absence of this translation, it should be obvious that a Zeus machine could verify that each \( \alpha \) holds of each of the sentences corresponding to the natural numbers in the infinitary version of Yablo’s paradox given above.

Are there arguments in the literature for the view that we are (or encompass) Zeus machines and the like? Yes. For example: After presenting the relevant mathematical landscape, Kugel (1986) argues that human persons are trial-and-error machines. Bringsjord (1992) gives a sustained argument for the view that human persons are Zeus machines, and has also argued (1997) specifically that logicians who work with infinitary systems routinely and genuinely use the \( \omega \)-rule. Again, the claim isn’t that such arguments are sound, and that therefore some human persons, contra Priest, genuinely use the \( \omega \)-rule. The claim is a simple, undeniable one: if any of these arguments are sound, then we can really use the \( \omega \)-rule, and the infinitary reasoning we gave above would appear to be above reproach.

But are the arguments sound? One of us believes so— but clearly there is insufficient space to present and defend them here. Fortunately, we can avail ourselves of a shortcut: we can consider the intuitions that precede and give rise to the arguments in question. In taking this shortcut, we follow a distinction affirmed by Turing (1939):

Mathematical reasoning may be regarded rather schematically as the exercise of a combination of two faculties, which we may call intuition and ingenuity. The activity of the intuition consists in making spontaneous judgments which are not the result of conscious trains of reasoning … The exercise of ingenuity in mathematics consists in aiding the intuition through suitable arrangements of propositions, and perhaps geometrical figures or drawings. (Turing 1939: 214–215)

It seems to us that humans often make correct spontaneous judgments about logical/mathematical issues because they see\( m \) that certain propositions are true; moreover, it would appear that what is seen\( m \) is that every member in some denumerable sequence has some property \( \alpha \). (This is not to say that after such judgments are made, these humans don’t show great ingenuity in sometimes crafting corresponding finitistic proofs making no use of the \( \omega \)-rule.) Suppose we remind a student that \( \mathbb{N} \) denotes the natural numbers \( \{0, 1, 2, \ldots\} \). Isn’t there a sense in which this student may see\( m \) the natural numbers? If we tell the student that to 1 there corresponds the sentence (\( p_1 \)) ‘1 is a natural number less than all subsequent natural numbers,’ and to 2 there corresponds (\( p_2 \)) ‘2 is a natural number less than all subsequent natural numbers,’ and so on \( ad infinitum \) in this fashion, is there not a definite, unobjectionable sense in which the student sees the sequence? Given a truth predicate \( T \), the student will doubtless agree that \( Tp_1, Tp_2, Tp_3, \ldots \); and the student, on the strength of this sequence, may well agree that \( Tp_n \), for all \( n \in \mathbb{N} \). At this point, the student may indeed be encouraged to craft a proof of this general conclusion, and the proof may well make use of a variable ranging over \( \mathbb{N} \) (and, say, mathematical induction). But nonetheless before the proof it seems that the student’s “mental eye” has seen what is then usually dutifully communicated in finitary terms. On the other hand, if the student produces a proof expressed in some infinitary system, even his or her ingenuity (to use Turing’s term) may capitalize on the \( \omega \)-rule.\(^7\)

\(^6\)In truth, only Bringsjord is convinced that the arguments are sound. Van Heuveln is agnostic.

\(^7\)This is perhaps as good a spot as any to note that infinitary logic has been seen by many as simply a natural
Perhaps the proof of our position is in the pudding. Once we saw the infinitary version of Yablo’s paradox (as set out above), we found a way to express at least the gist of this paradox in finitary terms. Specifically, we encoded the core conception as
\[
\forall x (Px \leftrightarrow \forall y (y > x \rightarrow \neg Px)),
\]
where the universe of discourse is \( N \), and we proceeded to derive a contradiction from this formula and basic facts about the natural numbers, viz., \( \forall x \exists y y > x \) and \( \forall x \forall y \forall z ((x < y \land y < z) \rightarrow x < z) \). In addition, once you see Yablo’s paradox well enough to go à la Turing’s distinction from intuition to standard first-order ingenuity, it isn’t hard to see the paradox of the gods (or new Zeno) — so that you can produce a parallel finitary proof.9

Rensselaer Polytechnic Institute
Troy, NY, USA
selmer@rpi.edu
heuweb@rpi.edu

way to surmount Gödel’s incompleteness results (Barwise 1980), and that such thinkers don’t view the background logic used to reason about infinitary logics to necessarily be finitary, despite Forster’s apparent belief that a finitary metalanguage is invariably used to reason over infinitary systems. Even a glance at formal reasoning over infinitary systems (e.g., see the works on infinitary logic cited in footnote 1) will reveal that the metalanguage in question makes use of the very same constructions as those that distinguish infinitary systems from finitary ones.

8 Even efficient versions of the proof, if carried out in natural deduction systems (e.g., Barwise and Etchemendy’s \( \mathcal{F} \)), can be over 45 lines long. Using hyper-resolution after clausifying, the proof can be done in 12 lines.

9 The paradox is given orginally by José Benardete (1964: 259), and introduced to Analysis readers in more formal form by Priest (1999). Subsequent analysis is offered by Yablo (2000) and Laraudogoitia (2000). Oddly enough, Yablo (2000) himself has written about the paradox, but makes no explicit connection between it and the paradox that bears his name. (Yablo does put the relevant paper of his in the references, and he numbers the demons with natural numbers in an increasing manner.) For the finitary proof thatparallels our finitary version of Yablo’s paradox, consider a demon whose intention is to freeze anything as soon as it moves even the slightest bit out of place. In this world, consider an object \( a \) whose potential movement would be in a certain direction, i.e., \( a \) starts at 0 and moves in a certain positive direction. Limit the universe of discourse to \( \mathbb{R}^+ \), and invoke the following predicates: \( Rx: a \) reaches \( x \); \( Fx: a \) gets frozen at \( x \). One can then prove, with just a little bit of ingenuity, that the following quintet is inconsistent by reasoning that parallels our finitary version of Yablo.

1. \( \forall x \exists y y < x \)
2. \( \forall x ((\exists y (Fy \land y < x)) \rightarrow \neg Rx) \) This is (2) from (Priest 1999).
3. \( \forall x ((\neg \exists y (Fy \land y < x)) \rightarrow Rx) \) This is (3) from (Priest 1999).
4. \( \forall x (Rx \leftrightarrow Fx) \) This is a modified version of (5) from (Priest 1999).
5. \( \forall x \forall y \forall z ((x < y \land y < z) \rightarrow x < z) \) Transitivity again.

For additional examples of “shortcut intuitions” that would seem to involve seeingm, see (Penrose 1994). Please note that these examples can be effective in the context of our case for seeingm, despite the real possibility that Penrose’s Gödelian arguments against computationalism fail. Indeed, one of us takes himself to have formally refuted these arguments in (Bringsjord & Xiao 2000).

We are indebted to an anonymous referee for trenchant objections, and to Steven Hay for catalyzing our search for formal equivalence between the Paradox of the Gods and finitary approximations of Yablo’s Paradox.
References


