Eduardo Alejandro Barrio

Theories of Truth without Standard Models and Yablo's Sequences

Abstract. The aim of this paper is to show that it's not a good idea to have a theory of truth that is consistent but ω-inconsistent. In order to bring out this point, it is useful to consider a particular case: Yablo's Paradox. In theories of truth without standard models, the introduction of the truth-predicate to a first order theory does not maintain the standard ontology. Firstly, I exhibit some conceptual problems that follow from so introducing it. Secondly, I show that in second order theories with standard semantics the same procedure yields a theory that doesn't have models. So, while having an ω-inconsistent theory is a bad thing, having an unsatisfiable theory of truth is actually worse. This casts doubts on whether the predicate in question is, after all, a truth-predicate for that language. Finally, I present some alternatives to prove an inconsistency adding plausible principles to certain theories of truth.

Keywords: Yablo's Paradox, non-standard models, ω-inconsistency, theories of truth.

The initial formulation of Yablo's Paradox[1] consists in an infinite set of sentences that is linearly ordered. Each of them claims that all sentences occurring later in the series are not true. At least at a superficial level, the series doesn't seem to involve any kind of self-reference.[2] According to Yablo, the set of sentences would be incapable of having a model and, therefore, it would be unsatisfiable. Of course, this result is controversial. It has been criticized by Priest and Ketland.[3] In particular, starting from a formulation of the series expressed in the language of arithmetic, Ketland shows that it is possible to find a non-standard model for Yablo's sequences. In this paper, I argue that Yablo's sequences introduces new boundaries

[1] [23] and [25].

[2] Far from a general consensus on this, there are many who claim that Yablo didn't manage to show that there is not some kind of circularity involved in his sequence. Mainly, Priest and Beall belong to this group and Bueno, Colyvan, Leitgeb, Sorensen and Yablo have maintain that Yablo's list generates a semantic paradox without circularity. See [1, 2, 13, 16, 20] and [25].

[3] See [16] and [11].

Presented by **Jacek Malinowski**; *Received* January 26, 2009

Studia Logica (2010) 96: 375–391
DOI: 10.1007/s11225-010-9289-8

to expressive capabilities of certain axiomatic theories of truth. Specifically, I show that adding Yablo's sequences of sentences to a certain theory of truth would generate expressive disorders. Adding Yablo's list of sentences and the *Local Yablo Disquotational Scheme* to first order arithmetic produces a theory of truth that is ω-inconsistent, but not inconsistent. The aim of this paper is to show that ω-inconsistency lead to unwanted results. So it's not good to have a theory of truth that is consistent but ω-inconsistent. In theories of truth without standard models, the introduction of the truth-predicate to a first order theory does not maintain the standard ontology. Firstly, I present a number of conceptual problems that follow from such an introduction. Secondly, I show that in second order theories with standard semantic the same introduction produces a theory that doesn't have a model. So, if an ω-inconsistent theory of truth is bad, an unsatisfiable theory is *really* bad. Finally, I present some alternatives to prove an inconsistency adding plausible principles to certain theories of truth. This casts doubts on whether the predicate in question is, after all, a truth-predicate for that language.

I.-

Yablo's paradox involves an infinite sequence of sentences Y_k, each of them stating that all the sentences occurring later in the series are not true.

(Y_0) For all k > 0, Y_k is not true.
(Y_1) For all k > 1, Y_k is not true.
(Y_2) For all k > 2, Y_k is not true.
(Y_3) For all k > 3, Y_k is not true.
And so on.

Informally, Yablo describes what happens with his sequence in the following way[4]: suppose, for *reductio*, that some particular sentence in the sequence is indeed true. For example, let's take Y_1 as true. Y_1 says that for all k > 1, Y_k is not true. Accordingly, we may conclude the following:

Y_2 is not true,
and
For every k > 2, Y_k is not true.

The latter, however, entails that what Y_2 says is in fact the case. This contradicts the first assumption: Y_2 turns out to be true. Thus, Y_1 must

[4] [23].

be false after all. The proof can be generalized, then, for each number k, we can prove that Y_k is false. Thus, for all k > 1, Y_k is not true. Therefore, Y_1 is true, which we have just shown to be impossible.

II.-

As is well known that from the previous informal proof Yablo intends to extract important consequences pertaining to the concept of *truth*. Since the proof doesn't seem to appeal to self-referential expressions, Yablo's infinite sequence of sentences seems to show that the aforementioned feature is not necessary for the existence of a semantic paradox. No doubt self-referentiality is hard to characterize, and its links to the concept of *circularity* are not at all helpful to this task.[5] The presence of self-referential features within the language has usually been assesed using techniques related to the arithmetization of formal languages. Let $\mathcal{L}_{\mathcal{PA}}$ be the language of Peano Arithmetic. $\mathcal{L}_{\mathcal{PA}+}$ is $\mathcal{L}_{\mathcal{PA}}$ augmented with a new unary predicate *Tr*. It's useful to assume that $\mathcal{L}_{\mathcal{PA}}$ includes a finite set of function-symbols for primitive recursive functions; particularly, we assume that this language has symbols for basic syntactic operations on $\mathcal{L}_{\mathcal{PA}+}$ in some fixed Gdel numbering. If α is a sentence of $\mathcal{L}_{\mathcal{PA}+}$, then $\ulcorner\alpha\urcorner$ denotes a term for the Gdel number of α. If α is a formula with one free variable, the term $\ulcorner\alpha(\dot{x})\urcorner$ is built from symbols for primitive recursive syntactic operations. $\ulcorner\alpha(\dot{x})\urcorner$ is an *open* function term, with x free, meaning 'the result of substituting the numeral of the number x for all free variables in the formula α'. More exactly, $\ulcorner\alpha(\dot{x})\urcorner$ can be defined as *sub(num(x))*, $\ulcorner\alpha(x)\urcorner$, where the function term *sub(x, y)* means 'the result of substituting x for all free variables in y' and *num(x)* means 'the numeral of x'. In contrast, note that $\ulcorner\alpha(x)\urcorner$ is a *closed* quotation term, in which the variable x is *not* free. This notation can then be extended to multiple variables. In $\ulcorner\phi(\dot{x_1}), \ldots, (\dot{x_n})\urcorner$, the dots above the variables indicate that these symbols are bound from outside in the usual way, where a function replaces the variables by the corresponding numerals. Dot notation was proposed by Feferman to allow quantification into formulae containing quotation terms.[6]

Let $\mathcal{PA}^1$ be Peano Arithmetic formulated in $\mathcal{L}_{\mathcal{PA}+}$. We could use $\mathcal{L}_{\mathcal{PA}+}$ in order to express true in $\mathcal{L}_{\mathcal{PA}}$. In this case, the predicate Tr, which is not in $\mathcal{PA}^1$, could be part of a consistent theory of truth in $\mathcal{PA}^1$.[7] In this

[5]Moreover, some authors think that self-reference cannot be explained at all. See [13].

[6][3, p.13].

[7][14] has shown that the strategy of adopting some kind of limitation to the T-schema

regard, we could add the set of all instances of the scheme '$\mathrm{Tr}(\ulcorner\phi\urcorner) \leftrightarrow \phi$' as an axiom scheme for all sentences of $\mathcal{L}_{\mathcal{PA}}$. Therefore, '$\phi$' cannot contain the symbol 'Tr' and the diagonalization theorem does not apply to that predicate because it applies only to formulae of $\mathcal{PA}^1$. Usually, the theory $\mathcal{PA}^1 \cup \{\mathrm{Tr}(\ulcorner\phi\urcorner) \leftrightarrow \phi$, for every sentence ϕ of $\mathcal{L}_{\mathcal{PA}}\}$ is called $\mathcal{TB}$ or $\mathcal{TD}$. This theory adds to $\mathcal{PA}^1$ all local Tarskian biconditionals.[8] It is easy to show that a simple modification of $\mathcal{TB}$ is enough to express Yablo's sequence of sentences inside $\mathcal{L}_{\mathcal{PA}+}$.[9] Following Ketland,[10] we could describe Yablo's list

---

suffers from different problems. Mainly, there are uncountably many different maximally consistent sets of instances of the T-scheme. And some of them are too complex (they are not recursively enumerable). Then, we need a method for sorting out the good instances from the bad instances of T schema. Even worse, as [8] shows, "given an arithmetical sentence (i.e., a sentence not containing Tr) that can neither be proved or disproved in $\mathcal{PA}$, one can find a consistent T-sentence that decides this sentences. (...) This implies that many consistent sets of T-sentences prove false arithmetical statements." Maximally conservative sets of instances of the T schema could be a good option. If we adopt an axiomatic approach, we can add certain truth-theoretic axioms to $\mathcal{PA}$. Tarski's Theorem says that adding the full unrestricted T-schema as an axiom to $\mathcal{PA}$ produces an inconsistent theory.

[8]Two main reasons support the adoption of $\mathcal{TB}$ as an axiomatization of the notion of *true in* $\mathcal{PA}$. As Tarski has proved, $\mathcal{TB}$ satisfies convention-T and it is consistent with respect to $\mathcal{PA}$. And if one thinks that truth does not have any explanatory force, as the deflationist claim, the new axioms for the truth predicate should not allow us to prove any new theorems that don't already involve the truth predicate. $\mathcal{TB}$ is conservative with respect to $\mathcal{PA}$. The new axioms for truth are conservative if they do not imply any additional sentences (free of occurrences of the truth-predicate) that are not provable without the truth axioms. According to this theory, the notion of *truth* serves mainly to express infinite conjunctions. $\mathcal{TB}$ helps to clarify exactly which infinite conjunctions can be expressed with a truth predicate. Of course, $\mathcal{TB}$ is a highly incomplete theory, because it does not prove the law of excluded middle, that is, the sentence:

$$(\forall\phi\ (\mathrm{Tr}(\phi) \lor \mathrm{Tr}(\neg\phi))$$

where the quantifier $\forall\phi$ is restricted to sentences not containing Tr. See [8].

[9]This move seems plausible not only because often truth is taken as a predicate applied to Gödelian numbers, but also because of one would expect that adding axioms or principles to characterize the truth predicate within a consistent theory yields a new richer theory that preserves consistency.

[10]Of course, it could not be completely clear why Ketland's formulation is actually Yablo's Paradox. The differences are clear: the original version is formulated in natural language and not in $\mathcal{L}_{\mathcal{PA}+}$. It doesn't need any particular notation in order to allow quantification into formulae containing quotation term. Nevertheless, and beyond the formalization method adopted, both Ketland and Leitgeb have connected Yablo's Paradox to theories that, being ω-inconsistent, have not standard models. For this reason, even when Ketland's proposal can be doubtful, my point is aimed against the acceptability of theories of truth lacking standard models.

of sentences as a sequence of biconditionals. Of course, what sustains such a movement is the idea of representing the application of the T-schema to any sentence of type Y_n. In order to express this sequence in $\mathcal{L}_{\mathcal{PA}+}$, we need to add an unary predicate Y(x). The open $\mathcal{L}_{\mathcal{PA}+}$-formula Y(x) might be called the *Yablo predicate*. The Yablo sentences are the sentences Y(1), Y(2), etc. Y(n) has the following intended interpretation: for all x > n, Y(x) is not true. The list of Yablo biconditionals are all instances of the set $\{Y(n) \leftrightarrow \forall k > n, \neg\, Tr(\ulcorner Y(\dot{k})\urcorner)\text{: } n \in \omega\}$. Finally, we could consider the *Local Yablo Disquotation Scheme*:

$$Tr(\ulcorner Y(n)\urcorner) \leftrightarrow Y(n), \text{ with } n \in \omega$$

One could expect that adding the list of Yablo biconditionals and the Local Yablo Disquotation Scheme to $\mathcal{PA}^1$ yields an inconsistency. Let $\mathcal{T}_{\mathcal{YA}}$ be the axiomatic theory of truth $\mathcal{PA}^1\cup$ the Local Yablo Disquotation Scheme (LYD) $\cup$ $\{Y(n) \leftrightarrow \forall k > n, \neg Tr(\ulcorner Y(\dot{k})\urcorner)\text{: } n \in \omega\}$.[11] As Ketland shows,[12] this theory is ω-inconsistent.[13] In order to appreciate the point with more detail, let A(k) be = Tr(Y(k)). Then, assume that:

1. $\mathcal{T}_{\mathcal{YA}} \vdash Tr(Y(1))$
2. $\mathcal{T}_{\mathcal{YA}} \vdash Y(1)$ by 1 and the LYD Scheme
3. $\mathcal{T}_{\mathcal{YA}} \vdash \forall k > 1, \neg\, Tr(\ulcorner Y(\dot{k})\urcorner)$ by 2 and Yablo's biconditionals

---

[11] $\mathcal{T}_{\mathcal{YA}}$ is the weakest theory of truth that can express the sentences of Yablo's sequence. Of course, one could add the Local Yablo Disquotational Scheme to $\mathcal{T}(\mathcal{PA})$. This theory is obtained when one adds to $\mathcal{PA}$ all induction axioms involving the truth predicate $\mathcal{PA}\cup$ {there is a full inductive satisfaction class}. $\mathcal{T}(\mathcal{PA})$ is not conservative over its base theory $\mathcal{PA}$. See [6]. In both cases, the result would also be ω-inconsistent. But, $\mathcal{T}(\mathcal{PA})$ is ω-consistent and induction axioms of $\mathcal{T}(\mathcal{PA})$ don't play any role in Yablo's sequence. See [10].

[12] [11, p.297].

[13] Formally, if a set of formulas is ω-inconsistent, it can be proved that even when each natural number fulfils the condition A(x), there is a number that doesn't. Of course, we are unable to prove that there is a specific number that doesn't fulfil the condition, but we can prove that there is one that doesn't fulfil it. Moreover, Gödel's results imply that $\mathcal{PA}^1$ is ω-incomplete, if $\mathcal{PA}^1$ is consistent: there are cases where it can be proved, case by case, that each number satisfies some condition A(x) but it can't be proved that all numbers satisfy A(x). Assuming that $\mathcal{PA}^1$ is consistent, Gödel's sentence can't be proved. Nevertheless, compare with a theory that results ω-inconsistent. It seems that ω-incompleteness in a theory of arithmetic is a regrettable weakness, but ω-inconsistency is a very bad news (nor as bad as outright inconsistency, of course, but still bad enough.) An ω-inconsistent theory can prove each of A($\underline{n}$) and yet also prove $\neg\forall$ x A(x) is just not going to be an acceptable candidate for expressing arithmetic. And if $\mathcal{T}_{\mathcal{YA}}$ is ω-inconsistent then $\mathcal{T}_{\mathcal{YA}}$'s axioms can't all be true on a standard arithmetic interpretation.

4. $\mathcal{T}_{\mathcal{YA}} \vdash \forall \mathrm{k} > 2, \neg \mathrm{Tr}(\ulcorner \mathrm{Y}(\dot{k})\urcorner)$ by $\mathcal{PA}^1$
5. $\mathcal{T}_{\mathcal{YA}} \vdash \neg \mathrm{Tr}(\mathrm{Y}(2))$ by 2
6. $\mathcal{T}_{\mathcal{YA}} \vdash \mathrm{Tr}(\mathrm{Y}(2))$ by 3
7. $\mathcal{T}_{\mathcal{YA}} \vdash \perp$ by 5 and 6 and Logic
8. $\mathcal{T}_{\mathcal{YA}} \vdash \mathrm{Tr}(\mathrm{Y}(1))$ Logic
9. $\mathcal{T}_{\mathcal{YA}} \vdash \neg \mathrm{Tr}(\mathrm{Y}(2))$, $\mathcal{T}_{\mathcal{YA}1} \vdash \neg\mathrm{Tr}(\mathrm{Y}(3))$, $\mathcal{T}_{\mathcal{YA}1} \vdash \neg\mathrm{Tr}(\mathrm{Y}(4))$, etc
10. $\mathcal{T}_{\mathcal{YA}} \vdash (\mathrm{Y}(1))$ by 8 and LYD Scheme
11. $\mathcal{T}_{\mathcal{YA}} \vdash \exists \mathrm{k} > 1, \mathrm{Tr}(\ulcorner \mathrm{Y}(\dot{k})\urcorner)$ by 10 and Yablo's biconditionals
12. $\mathcal{T}_{\mathcal{YA}} \vdash \exists \mathrm{k}, \mathrm{Tr}(\mathrm{Y}(\dot{k}))$ by 11 and $\mathcal{PA}^1$

9 and 12 imply that $\mathcal{T}_{\mathcal{YA}}$ is ω-inconsistent.[14]

This result implies that $\mathcal{T}_{\mathcal{YA}}$ has not standard model. That is, there is no model whose domain is the set of natural numbers in which Yablo's list could acquire a consistent interpretation. This entails that no expansion of a standard model $\mathfrak{N}$ of $\mathcal{PA}$ satisfies $\mathcal{T}_{\mathcal{YA}}$. However, $\mathcal{T}_{\mathcal{YA}}$ has a non-standard model. It is worth recalling that a standard model for arithmetic is any model of the set of all sentences of the language of arithmetic that are true according to the standard interpretation of arithmetic. The existence of non-standard models for arithmetic, models that are not isomorphic to the standard interpretation, is a direct consequence of the application of *the Compactness Theorem* for first order languages. It appeals to an extended model to which not only standard natural numbers belong but also non-standard numbers, and then shows that the complete list, numbered with the standard natural numbers, acquires, in this model, a consistent assignment of truth values. Obviously, the key point is that all the new non-standard elements are larger than the standard numbers that helped us number the sentences of Yablo's sequence. Then, the formalization of Yablo's sequence inside $\mathcal{T}_{\mathcal{YA}}$, as part of the language of first order arithmetic, is satisfiable by at least one non-standard model of arithmetic;[15] and for this reason, since

---

[14]McGee's Theorem shows the conditions that a truth theory has to satisfy to be ω-inconsistent. [14].

[15]Peano arithmetic theory formalized in first order language is not categorical. That is, it can have models different from the standard (the one that has as domain the set of numbers {0, 1, 2 ...}). These other models (non-standard models) have domains, including the standard one, but add another list not isomorphic of elements {0, 1, 2, ..., -2*, -1*, 0*, 1*, 2* ...}. Each one of the elements in this list is larger than each of the standard numbers. Thus, new non-standard models can be obtained by adding new lists to the domain of interpretation. See [9].

Yablo's set of sentences formulated in the language of first order arithmetic has a model, it is formally consistent.[16]

But, let me emphasize: although the sequence formalized in this way is not inconsistent, it turns out to be ω-inconsistent.[17] This point is disconcerting: against what we would have thought, it is possible to find a model for $\mathcal{T}_{\mathcal{YA}}$, after all. However, this model cannot have as its domain the set of the standard natural numbers. For this reason, even though Yablo's sequence is consistent, it has not at the same time standard model. Thus, even though the theory that includes the list of Yablo's sentences is not strictly paradoxical (since it has non-standard models), it cannot be interpreted using a structure in which the numbers appearing in each of the biconditionals are standard natural numbers. Because of this, even if the sequence is consistent, it turns out to be ω-inconsistent. Thus, according to Ketland, Yablo's Paradox is not strictly a paradox but actually an ω-paradox.

III.-

Tarski's Theorem limits our capacities of expressing all instances of the T-schema within $\mathcal{PA}^1$: if the set of all the T-sentences is added to $\mathcal{PA}^1$, the resulting theory will be inconsistent and, therefore, it will lack a model. It is usually interpreted to mean that $\mathcal{PA}^1$ cannot express its own truth predicate. For this reason, when we do obtain an inconsistency, we tend to consider that the set of these instances becomes inexpressible within $\mathcal{PA}^1$. But, adding the Local Yablo Disquotation Scheme to $\mathcal{PA}^1$ produces a consistent first order theory, which is, nevertheless, ω-inconsistent. That is, $\mathcal{T}_{\mathcal{YA}}$ does have a model: if Yablo's sequence is expressed within first order arithmetic, it is satisfied by a structure whose domain includes, besides standard numbers, non-standard natural numbers. In what follows I want to emphasize that even if the sequence of Yablo sentences can have non-standard models, it will never have a standard model anyway. From my point of view, this result has significant consequences regarding our capacity to express the concept of *truth*.[18] As with Tarski's theorem, it is a direct consequence of Ketland's

---

[16] Other proof of consistency for $\mathcal{T}_{\mathcal{YA}}$ is quite simple: if there were a negation inconsistency, then, by compactness, the inconsistency would be confinable to a finite set of sentences. Nevertheless, no finite subset of Yablo biconditionals are inconsistent.

[17] Vann McGee describes some conditions that must be satisfied in order to produce an ω-inconsistency. Hannes Leitgeb generalizes these results to different constructions. See [14] and [12].

[18] There are several axiomatic theories of truth that lack standard models. In particular, [5] proves that Friedman and Sheard's theory $\mathcal{FS}$ is also ω-inconsistent. This result is very

result that by adding enough expressive resources to $\mathcal{PA}^1$ to talk about the infinite ordered sequences of sentences like Yablo's, we limit the capacity of expressing the truth within the resulting theory. And this happens not because the theory doesn't have a model and for this reason it is inconsistent, but because even if we were able to find models for it, the alleged truth predicate of the resulting theory would not be a good representation of the truth. In other words, from my perspective, the sequence of Yablo sentences shows us that a theory such as $\mathcal{T}_{\mathcal{YA}}$ is unable to express the truth predicate.

Every successful formal theory of truth must be supported by philosophical intuitions. If not, our attempted axiomatization of truth might not express a legitimate truth predicate. But in my opinion, $\mathcal{T}_{\mathcal{YA}}$ is not an acceptable truth theory $\mathcal{T}_{\mathcal{YA}}$ since it is not able to preserve the intuition that adding a theory of truth to some base theory having a model should not interfere with the intended ontology of that base-theory. Specifically, regarding $\mathcal{PA}^1$, the addition of Yablo's sequence wouldn't have to disturb the intended arithmetical ontology. Nevertheless, as we have seen, it does disturb it. This is mainly because $\mathcal{T}_{\mathcal{YA}}$ itself lacks a standard model. In this sense, $\mathcal{T}_{\mathcal{YA}}$ defines truth conditions for the formulas of $\mathcal{PA}^1$ that won't depend on $\mathcal{PA}^1$'s intended ontology.

Note that Ketland's proposed resolution of Yablo's Paradox implies giving up the possibility of gaining a better understanding of the underlying domain of the theory: the truth-free part of the theory *has ω-models* for arithmetic, but the theory of truth that enlarges that theory has no ω-models.[19] That is, $\mathcal{L}_{\mathcal{PA}}$ is a first order language enriched with enough expressions as to express arithmetic. According to its standard interpretation, formal numerals 0, 1, 2, 3 in the official language of $\mathcal{PA}^1$ refer to the numbers 0, 1, 2, 3 respectively. Similar considerations apply to predicates and function symbols expressing arithmetical concepts. And quantifiers will have the standard natural numbers as their scope. In this way, the standard interpretation warrants that expressions of arithmetic such as '3 + 0 = 3' or '∀x(x + 0 = x)' talk about natural numbers. Of course, there are no categoricity results forbidding the existence of non-standard models. But the important thing is that $\mathcal{PA}^1$'s axioms and theorems come out to be true in those structures whose domains are exclusively composed of natural numbers.

important because $\mathcal{FS}$ is usually considered a complete and consistent system. Of course, the objections that I describe in this paper also apply to $\mathcal{FS}$.

[19] An ω-models is any model for $\mathcal{PA}$ that is isomorphic to the intended model of arithmetic $\langle \omega, +, \bullet, 0, 1, < \rangle$. One important difference between ω-models and non-ω-models is that non-ω-models don't automatically satisfy full induction.

Now, part of the reasons why we have to introduce a truth-predicate applicable to $\mathcal{PA}^1$'s axioms and theorems is that we wish that '∀x(x + 0 = x)' would come out true exactly when this formula talks about those standard numbers. In other words, the enlargement of $\mathcal{PA}^1$ that allows the application of an alleged truth-predicate to $\mathcal{PA}^1$'s formulas, obtaining in turn formulas such as '*Tr*(∀x(x + 0 = x))', shouldn't alter $\mathcal{PA}^1$'s intended ontology. Formally, in order to preserve this intuition we need that the enlarged theory be ω-conservative: if $\mathcal{PA}^1$ *has ω-models*, the extension of $\mathcal{PA}^1$ that includes occurrences of an alleged truth-predicate *must also have ω-models*. Notwithstanding, as we have seen, Yablo's sequences along with arithmetic, do not comply with this result. It can indeed have a model, but its models have, besides standard elements, non-standard ones. For this reason, the introduction of the alleged truth-predicate doesn't maintain the standard ontology and therefore, the expression ends up being unable to express legitimate truth.

In any case, given that we want formal arithmetic to have axioms which are true on a standard interpretation, we must want ω-consistent arithmetic. And given that we believe that arithmetic is sound on its standard interpretation, we are committed to thinking that it is ω-consistent. So, adding certain truth-theoretic axioms to $\mathcal{PA}^1$ must not yield an ω-inconsistent theory.

The fact that the theory that includes the Yablo's biconditionals doesn't have any ω-models means that although the sentence '*Tr*(∀x (x + 0 = x))' is a theorem of the theory, its truth cannot be seen as concern with standard natural numbers. Given that the theory including them does not have a model whose domain is that of standard natural numbers, the sentences of the theory cannot be seen as truths about the theory's intended ontology. Still, the very distinction between the standard and non-standard models seems to assume that semantic notions, and particularly the language's truth-predicate have a standard interpretation. For this reason, our ensuring that there is no standard model for Yablo's sequence is conceptually sufficient for ensuring that a truth predicate that is part of the expressive resources of a first order language that allows to express infinite ordered series of sentences does not represent a legitimate concept of truth. What ω-consistency guarantees is that arithmetic truth depends on the ontology of arithmetic. However, since they end up being ω-inconsistent, Yablo's sequences represent an alleged truth-predicate whose application conditions cannot depend on natural numbers. Otherwise, whatever the principles are that allow us to establish that a monadic predicate expresses the truth-predicate, they should guarantee that the mentioned predicate applies to certain expressions of the language, assuring that the result offers *the* right extension for the concept of *truth*. But if, following Ketland, we concede that the components of Yablo's sequence are interpreted by means of non-standard models, the interpretation of the alleged truth-predicate does not seem to offer an appropriate interpretation for that language. In any case, it seems to offer a characterization of a *non-standard* concept of truth.

In sum, ω-consistency seems to be a highly desirable feature of a theory of truth. Ketland's result proves that theory $\mathcal{T}_{\mathcal{Y}\mathcal{A}}$ doesn't comply with this feature, and hence is unable to correctly express the semantic properties of the sentences of $\mathcal{PA}^1$. If a theory of truth is not ω-consistent, it may not be interpreted as talking about the intended ontology of the theory to which it applies. Thus, the enlargement of $\mathcal{PA}^1$ with Yablo's sequence should have an ω-model. Only in this way we can have certain warrants that $\mathcal{T}_{\mathcal{Y}\mathcal{A}}$ expresses a legitimate truth predicate. No monadic predicate that we add to $\mathcal{PA}^1$ will express legitimate truth if its introduction to the language of arithmetic produces in turn *a dramatic deviation in the theory's intended ontology*: in order to be able to express the concept of *arithmetic truth*, $\mathcal{T}_{\mathcal{Y}\mathcal{A}}$ has to abandon the possibility of speaking about standard natural numbers. For this reason, ω-consistency is an additional requirement that should be satisfied every time we are concerned with expressing truth: not only do we want the addition of Yablo's sequence to $\mathcal{PA}^1$ to result in a consistent theory, but we also want it to result be capable of conserving the ontology of the intended interpretation of $\mathcal{PA}^1$, for it seems plausible to maintain that no theory of truth should imply a substantive answer to what numbers are or to what the ontology of arithmetic is.

However, the proposal to treat ω-consistency as an adequacy condition for any theory intending to express legitimate truth is not out of discussion. There are several recent theories of truth that lack a standard model.[20] In this direction, Michael Sheard argues that "(the) fascinating discovery that some consistent axiomatic theories of truth are in fact ω-inconsistent does

[20]In the last few years several constructions have emerged that, as Yablo's series, are consistent but ω-inconsistent. All of them lead to results of non-standardness rather than to inconsistency. In this line of thought, McGee has presented an infinite Liar-sentence. This sentence asserts that not every iterated application of the truth predicate to itself will be true. There is a non-standard model for this sentence, even when, of course, it lacks a standard model. Visser has presented a construction analogous to McGee's that, unlike the latter, involves a non-wellfounded hierarchy of languages, each level with their respective truth predicates. Leitgeb, instead, has formulated an ω-inconsistent construction assuming the existence of beings with infinite capacities. It can be shown that Uzquiano's paradox on the denotation of certain definite descriptions within $\mathcal{PA}$ and that of the Gods blocking the road to a man also generate ω-inconsistencies. See [12, 14, 21, 22] and [24].

not present a significant impediment to the effective use of those theories".[21] Volker Halbach and Leon Horsten support the same point of view.[22] They show that there are some truth theories that are ω-inconsistent. In the first place, because theories with this characteristic are arithmetically sound, that is, this type of theories doesn't prove any false sentence of arithmetic. So, the ω-inconsistency concerns only that part of the theory that deals with the truth predicate.[23] But, my point is precisely that ω-inconsistency implies that the theory that includes the truth predicate is not conservative in ω-models. The lack of attractiveness is not due to the fact that we can prove false sentences of arithmetic, but that of we cannot preserve arithmetic's intended ontology.

To this, Halbach and Horsten could answer that the idea that such theories don't have nice models is at least questionable. "We accept set theory although we cannot prove that there is any nice model for set theory. Because of Gdel's second incompleteness theorem we cannot even prove that set theory has any model."[24] Halbach and Horsten agree that there are some differences between the case of set theory and the case in which we get a truth theory without an ω-model. But, they claim that set theory does not reject the existence of a strong inaccessible cardinal number, whereas in the case of this type of theory of truth, ω-inconsistency refutes the existence of a ω-model. So, contrary to what I think, Halbach and Horsten don't believe that these results make ω-inconsistent theories unacceptable.[25] They claim that the semantic functioning of these theories is very natural, given that a theory with a finite number of Yablo sentences possesses a nice standard model.[26] Since we can use only finite resources in any proof, at any step of our reasoning we will have a nice model. They seem to think that in the case of Yablo's sequence the ω-inconsistency just reflects the fact that there is no nice limit model at level ω. But, at the same time, the consistency of the sequence ensures that nothing is wrong with the theory.

Nevertheless, according to my view, there seems to be a difference between the impossibility of proving the existence of a nice model, and even between the impossibility of proving the existence of a model (as it happens in

---

[21] [18, p. 179].

[22] [7].

[23] [7, p. 213].

[24] [7, p. 214].

[25] I have discussed these ideas with Volker Halbach during a visit to Oxford. I thank him for valuable comments and suggestions.

[26] [7, p. 214].

the case of the set theory[27]), and the possibility of proving the non-existence of a nice model (as it happens in the case of truth theories that lack standard models). The last point seems to show that, unlike what happens in set theory, truth theories of the type we have been examining don't express legitimate truth, for we have enough proofs that show that there are no ω-models in which the theory ends up being truth.

To sum up, I take it that what Yablo's paradox enables us to draw the following morals: a theory of truth not only has to be satisfiable: it also has to be ω-consistent. Otherwise, if it wasn't, the theory wouldn't have an ω-model and in that case, it wouldn't have to be interpreted as referring to natural numbers but, instead, to non-standard numbers. In other words: the ω-inconsistency of a theory implies that its models are non-standard, and in the case the theory tries to express the truth predicate, its lack of standard models prevents the expressed truth from representing our intuition according to which the truth of $\mathcal{PA}^1$ should depend on standard natural numbers. For that reason, the fact that Yablo's sequence of sentences is ω-inconsistent should be enough to show a new kind of expressive incapacity: the one of representing, within a language, a truth predicate that establishes a close link between truth and the standard ontology. Because of this, the existence of non-standard models, even when it avoids inconsistency, generates a new expressive limitation related to truth: a language of arithmetic that is able to express Yablo's sequence does not have enough expressive power to have its own (standard) truth predicate.

IV.-

In fact, what we get when we formalize Yablo's sequence in a first order language is an ω-inconsistent theory of truth. But we should note here that the sequence turns out to be unsatisfiable under a formalization that includes higher-order resources with standard semantic. Dedekind shows

[27] However, McGee has recently shown that the axioms of second-order Zermelo-Fraenkel set theory plus choice with urelements (ZFCU), plus the axiom that urelements form a set, are able to characterize the structure of the universe of pure sets up to isomorphism. This result has to assume unrestricted quantification. Then, any two models of ZFCU+ the Urelement Set Axiom (in which our quantifiers take unrestricted range over all the objects there are) have isomorphic pure sets. Of couse, McGee's point doesn't tell us if there are seven inaccessible cardinals or if there are more or fewer, but it does tell us that the sentence of set theory that states that there are seven inaccesible cardinals is either true in all models of second order of ZFCU+ the Urelement Set Axiom in which our quantifiers take unrestricted range or false in all such models. In any case, the assumptions of the result is very controversial and I don't need it in order to defend my position. See [15].

that all models of second order arithmetic are isomorphic (given standard semantic).[28] Of course, it is crucial in order to set up this result that second order quantifiers are required to run over the full collection of subsets of domain. In this case, the existence of non-standard models is avoided.[29] For which, for each of Ketland's biconditionals, there is a corresponding element of the domain (a standard natural number). There is no possibility of there being a domain with an element larger than all standard natural numbers serving as a model for the sequence. Arithmetic, as formulated with second-order language, is categorical: any two models of either theory are isomorphic. Thus, adding the Local Yablo Disquotation Scheme to second-order Peano arithmetic ($\mathcal{PA}^2$) produces an unsatisfiable theory.[30] Let $\mathcal{T}^2_{\mathcal{YA}}$ be the axiomatic theory of truth $\mathcal{PA}^2 \cup$ the Local Yablo Disquotation Scheme $\cup$ $\{\mathrm{Y(n)} \leftrightarrow \mathrm{k} > \mathrm{n}, \neg\mathrm{Tr}(\ulcorner \mathrm{Y}(\dot{k})\urcorner)\text{: } \mathrm{n} \in \omega\}$. Then, $\mathcal{T}^2_{\mathcal{YA}}$ don't have a model. Obviously, this is a really decisive result. So, when one adds the Local Yablo Disquotation Scheme to first order arithmetic one produces a theory that has nonstandard models. But, when one adds the same scheme to $\mathcal{PA}^2$ one produces a theory that is unsatisfiable. Because of $\omega-$inconsistency, $\mathcal{T}_{\mathcal{YA}}$ has problems. But adding the Local Yablo Disquotation Scheme to a higher order theory with standard semantic is definitely a bad idea.

Obviously, this result doesn't show that Yablo's sequence formalized within $\mathcal{PA}^2$ would be syntactically inconsistent. Any proof must use only a finite subset of Ketland's bicondicionals. But, all such subsets are consistent. So, although second order arithmetic together with Yablo's sequence is unsatisfiable given standard semantics, it is not possible to prove an inconsistency within this theory. As Forster claims:[31] "The first thing to notice is that the proof of the [Yablo] paradox is infinitely long". The fact that there is a consistent formulation of the sequence is therefore not enough to

---

[28]For a modern proof, see [17, pp. 82–83].

[29]In standard semantics, relation variables range over the entire class of relations on the domain. A standard model of second order language is the same as a model of the corresponding first-order language, namely a domain and appropriate referents of the non-logical expressions. Then, when one specifies a domain, one thereby specifies the range of both the first-order variables and the second-order variables.

[30]$\mathcal{PA}^2$ is the strongest theory in a second-order language: it can quantify over arbitrary numerical sets. All models of $\mathcal{PA}^2$ are ω-models. Versions of second-order arithmetic which give up that requirement —theories which are built in two-sorted first-order languages— are not categorical. In this sense, $\mathcal{PA}^2$ is stronger than $\mathcal{ACA}_0$ (arithmetical comprehension axiom with restricted induction). This subsystem of second order arithmetic has non-ω-models. See [19].

[31][4].

prove that there is nothing wrong with the theory.[32] Of course, one could get an effect similar to the higher-order case by adopting the ω-rule or other infinitary resources capable of expressing Yablo's list as an infinite conjunction. In both cases, one could obtain a contradiction from the formulation of Yablo's sequence of sentences within a first order language. In this regard, Leitgeb points out[33] that if a theory formulated in the language of first order arithmetic is consistent in ω-logic, then it is ω-consistent. But $\mathcal{T}_{\mathcal{YA}}$is not ω-consistent. So, it is not consistent in ω-logic. Of course, for epistemological reasons, it is not obvious to accept infinitary resources to get an inconsistency from Yablo's sequence of biconditionals.

Another possibility to get an inconsistency from Yablo's sequence is to add the Uniform Yablo Disquotation Principle:

$$\forall \mathrm{x}(\mathrm{Y}(\mathrm{x}) \leftrightarrow \forall \mathrm{k} > \mathrm{x}, \neg \mathrm{Tr}(\ulcorner \mathrm{Y}(\dot{k}) \urcorner)$$

This principle is different from the Local Yablo Disquotational Scheme. Using that idea, one can semantically ascend and prove '$\forall \mathrm{x}(\mathrm{Tr}(\ulcorner \mathrm{Y}(\dot{x}) \urcorner) \leftrightarrow \forall \mathrm{k} > \mathrm{x}, \neg \mathrm{Tr}(\ulcorner \mathrm{Y}(\dot{k}) \urcorner)$'. But this formula implies an inconsistency. So, adding the Uniform Yablo Disquotation Principle to $\mathcal{PA}^2$ with standard semantics produces a theory that is inconsistent (not only unsatisfiable).

Finally, adding the Global Reflection Principle (GRP):

$$\forall \mathrm{x}(\mathrm{Sent}(\mathrm{x}) \ \& \ \mathrm{Bew}_{\mathcal{T}_{\mathcal{YA}}}(\mathrm{x}) \rightarrow \mathrm{Tr}(\mathrm{x}))$$

and the axiom:

$$\forall \mathrm{x}(\mathrm{Tr}(\ulcorner (\dot{x}) \urcorner) \leftrightarrow \mathrm{Tr}(\ulcorner \forall x \phi(x) \urcorner))$$

to $\mathcal{T}_{\mathcal{YA}}$ also produce an inconsistent theory. Reflection Principles are clearly schematic assertions of soundness — anything provable is true. As such, they imply consistency. So, if we added to $\mathcal{T}_{\mathcal{YA}}$, we would expect that the result be consistent.[34] Nevertheless, we get an inconsistent theory. The result focus attention on the role of the axiom $\forall \mathrm{x}(\mathrm{Tr}(\ulcorner (\dot{x}) \urcorner) \leftrightarrow \mathrm{Tr}(\ulcorner \forall x \phi(x) \urcorner))$.[35] This axiom says that a universally quantified sentence of

[32] I thank Øystein Linnebo, Agustn Rayo and Gabriel Uzquiano for stimulating discussion concerning this point.

[33] [12, p. 71].

[34] Obviously, Gödel's results imply that $\mathcal{T}_{\mathcal{YA}}$ can not prove its own Reflexion Principle. In this case, $\mathcal{T}_{\mathcal{YA}}$ might prove its own consistency.

[35] Several theories of truth have this axiom. $\mathcal{T}(\mathcal{PA})$— is one of them. This theory is conservative over $\mathcal{PA}^1$. Also, $\mathcal{T}(\mathcal{PA})$ and FS have this axiom: but both are not conservative over $\mathcal{PA}^1$.

the $\mathcal{L}_{\mathcal{PA}}$ is true if and only if all its numerical instances are true. This axiom is called (U-Inf) by Sheard.[36] (U-Inf) is crucial in the proof that follows:

PROOF SKETCH. [37] Firstly, by McGee's Theorem, $\mathcal{T}_{\mathcal{YA}}$ is ω-inconsistent. So, there is a formula A(x) of $\mathcal{L}_{\mathcal{YA}}$ with

$$1\ \mathcal{T}_{\mathcal{YA}} \vdash \neg\ \mathrm{A(n)}\ \text{for all}\ \mathrm{n} \in \omega.$$
$$\mathcal{T}_{\mathcal{YA}} \vdash \neg(\mathrm{A}(1)),\ \mathcal{T}_{\mathcal{YA}} \vdash \neg(\mathrm{A}(2)),\ \mathcal{T}_{\mathcal{YA}} \vdash \neg(\mathrm{A}(3)),\ \mathcal{T}_{\mathcal{YA}} \vdash \neg(\mathrm{A}(4)) \ldots$$

But also

$$2\ \mathcal{T}_{\mathcal{YA}} \vdash \exists \mathrm{x},\ \mathrm{A(x)}$$

The proofs of $\neg$A(1), $\neg$A(2), $\neg$A(3), $\neg$(A(4)), ... in $\mathcal{T}_{\mathcal{YA}}$ can be given in a straightforward uniform way. Then,

$$3\ \mathcal{T}_{\mathcal{YA}} \vdash \forall \mathrm{x}\ \mathrm{Bew}_{\mathcal{T}_{\mathcal{YA}}}(\ulcorner \neg \mathrm{A}(\dot{x})\urcorner)$$

But using the Global Reflection Principle (GRP), I conclude the following:

$$4\ \mathcal{T}_{\mathcal{YA}} \cup\ (\mathrm{GRP}) \vdash \forall \mathrm{x}\ \mathrm{Tr}(\ulcorner \neg \mathrm{A}(\dot{x})\urcorner)$$

Secondly, but using the left-to-right direction of (U-Inf), taking '$\neg$A' for 'ϕ', we get:

$$5\ \mathcal{T}_{\mathcal{YA}} \cup\ (\mathrm{GRP}) \cup (\text{U-Inf}) \vdash \mathrm{Tr}(\ulcorner \forall \mathrm{x}(\neg \mathrm{A(x)})\urcorner)$$

And using the Disquotation Principle in 5,

$$6\ \mathcal{T}_{\mathcal{YA}} \cup\ (\mathrm{GRP}) \cup (\text{U-Inf}) \vdash \forall \mathrm{x} \neg \mathrm{A(x)}$$

But, 2 and 6 imply

$$7\ \mathcal{T}_{\mathcal{YA}} \cup \{\mathrm{GRP}\} \cup (\text{U-Inf}) \vdash \bot^{38}$$

■

Of course, this result does not show that $\mathcal{T}_{\mathcal{YA}}$ is inconsistent. But assuming GRP and (U-Inf) because of being ω-inconsistent, $\mathcal{T}_{\mathcal{YA}}$ becomes inconsistent.

---

[36][18]

[37][14].

[38]I thank Volker Halbach for interesting discussion concerning this proof.

V.-

We have seen that there are several ways of introducing Yablo's sequence of sentences. All of them seem to have important consequences for the concept of *truth*. On the one hand, any formulation that enriches the resources of first order languages does not allow us to obtain models when expressing the sequence. On the other hand, its first order formulation is only capable of attaining a model that includes non-standard elements. Either case, it is possible to prove that when expressing a Yablo series within $\mathcal{T}_{\mathcal{YA}}$, a dramatic change in the intended ontology occurs. The resulting truth theory is ω-inconsistent, and hence it loses the capacity to talk about the standard natural numbers. So, in a theory of truth without standard models, arithmetic truth does not depend on the intended ontology of arithmetic. Moreover, in second order case, compared with the first order case, the result is even worse: higher-order resources with standard semantic avoid the existence of non-standard models. So, adding Yablo's sequence to second order arithmetic with standard semantic produce a theory of truth that doesn't have a model. If a theory of truth that be ω-inconsistent is a bad thing, having a unsatisfiable theory is really bad. A similar effect can be got by adopting ω-rule, other infinitary resources or certain intuitive principles of truth.

Acknowledgements. I am grateful to a large number of people for comments and discussion. I owe thanks to Eleonora Cresto, Volker Halbach, Ignacio Ojea, Lavinia Picollo and Gabriel Uzquiano as well as members of Logic Research Group at University of Buenos Aires. I am also heavily indebted to Øystein Linnebo and Agustín Rayo for detailed and helpful comments. Finally, I would like to acknowledge the suggestions of anonymous referee which allow me to significantly improve my paper.

References

[1] Beall, J. C., 'Is Yablo's Paradox Non-Circular', *Analysis*, 61 (2001), 176–187.

[2] Bueno, O., and M. Colyvan, 'Paradox without Satisfaction', *Analysis*, 63 (2003), 152–156.

[3] Feferman, S., 'Reflecting on Incompleteness', *Journal of Symbolic Logic*, 56 (1991), 1, 1–49.

[4] Forster, Th., 'The Significance of Yablo's Paradox without Self-Reference', 1996. Available at URL = <http://www.dpmms.cam.ac.uk/-tf>.

[5] Hablach, V., 'A System of Complete and Consistent Truth', *Notre Dame Journal of Formal Logic*, 35 (1994), 353–370.

[6] Hablach, V., 'Conservative Theories of Classical Truth', *Studia Logica*, 62 (1999), 353–370.

[7] Halbach, V., and L. Horsten, 'The Deflationist's Axioms for Truth', in J. C. Beall and B. Armour-Garb, (eds.), *Deflationism and Paradox*, Oxford University Press, Oxford, 2005, pp. 203–217.

[8] Halbach, V., 'Axiomatic Theories of Truth', in E. N. Zalta, (ed.), *Stanford Encyclopedia of Philosophy*, available at URL = <http://plato.stanford.edu/entries/truth-axiomatic>, 2007.

[9] Kaye, R., *Models of Peano Arithmetic*, Oxford University Press, Oxford, 1991.

[10] Ketland, J., 'Deflationism and Tarski's Paradise', *Mind*, 108 (1999), 429, 69–94.

[11] Ketland, J., 'Yablo's Paradox and ω-inconsistency', *Synthese*, 145 (2005), 295–302.

[12] Leitgeb, H., 'Theories of Truth which have no Standard Models', *Studia Logica*, 68 (2001), 69–87.

[13] Leitgeb, H., 'What is a Self-Referential Sentence? Critical Remarks on the Alleged (Non-)Circularity of Yablo's Paradox', *Logique et Analyse*, 177-178 (2002), 3–14.

[14] McGee, V., 'How Truthlike Can a Predicate Be? A Negative Result', *Journal of Philosophical Logic*, 14 (1985), 399–410.

[15] McGee, V., 'How We Learn Mathematical Language', *Philosophical Review*, 106 (1997), 35–68.

[16] Priest, G., 'Yablo's Paradox', *Analysis*, 57 (1997), 4, 236–242.

[17] Shapiro, S., *Foundations without Foundationalism: A case for Second-Order Logic*, Oxford Clarendon Press, Oxford, 1991.

[18] Sheard, M., 'Weak and Strong Theories of Truth', *Studia Logica*, 68 (2001), 89–101.

[19] Simpson, S., *Subsystems of Second Order Peano Arithmetic*, Cambridge University Press, Cambridge, 2009.

[20] Sorensen, R., 'Yablo's Paradox and Kindred Infinite Liars', *Mind*, 107 (1998), 137–155.

[21] Uzquiano, G., 'An Infinitary Paradox of Denotation', *Analysis*, 64 (2004), 138–131.

[22] Visser, A., 'Semnatics and the liar paradox', in D. Gabbay and F. Guenther, (eds.), *Handbook of Philosophical Logic*, Kluwer, Dordretch, 2004, pp. 149–204.

[23] Yablo, S., 'Paradox without Self-Reference', *Analysis*, 53 (1993), 251–252.

[24] Yablo, S., 'A Reply to New Zeno', *Analysis*, 60 (2000), 148–151.

[25] Yablo, S., 'Circularity and Paradox', in V. Hendricks, S. Pedersen, and T. Bolander, (eds.), *Self-reference*, CSLI Press, Stanford, 2006, pp. 139–157.

Eduardo Alejandro Barrio
Universidad de Buenos Aires
— Conicet — GAF
Buenos Aires, Argentina
eabarrio@gmail.com