The Yablo Paradox in Second Order: Consistency and Unsatisfiability

Lavinia Picollo

Draft! Not final version!

Abstract

Yablo [15], [16] introduces a new semi-formalized paradox, constituted by an infinite list of sentences, and informally gets a contradiction from it. However, it has been shown that, formalized in a first order language, Yablo's piece of reasoning is invalid, for it is impossible to derive *falsum* from the list, due mainly to the Compactness Theorem. This result casts doubts on the paradoxical character of the list of sentences. After identifying two usual senses in which an expression or set of expressions can be paradoxical, since second order languages are not compact, I explore the paradoxicality of Yablo's list within these languages. While non-paradoxical in the first sense, the second order version of the list is a paradox in our second sense. I conclude that this suffices for regarding Yablo's original list as paradoxical and his informal argument as valid.

1 Introduction

Infinitary paradoxes arose in the last decades to challenge the idea, almost undisputed until then, that the cause of antinomies lies on some sort of circularity or self-reference (a limit case of circularity) involved in the premises (core of the paradoxes) of the piece of reasoning that entails a contradiction. The first and most popular one was the Yablo Paradox, introduced by Yablo [15], [16]. All these antinomies consist in an infinite list of expressions, each of which says something only about other expressions below in the list, avoiding, *prima facie*, any kind of circularity. The Yablo Paradox is given by the following semi-formalized sentences:

(Y_0) For all $k > 0$, Y_k is untrue.

(Y_1) For all $k > 1$, Y_k is untrue.

(Y_2) For all $k > 2$, Y_k is untrue.

...

Informally, it is possible to establish the paradoxicality of the list by getting a contradiction from it. Suppose for contradiction that some Y_n is true. Since it says that the following are not, they are not. Then, Y_{n+1} is not true, and neither are all the sentences that stand below it in the list. But this is precisely what Y_{n+1} states. Then, Y_{n+1} must be true after all, which is absurd because we have just said it was not. Hence, every sentence Y_n in the sequence is false. In particular, all sentences subsequent to any given Y_n are false and, therefore, Y_n is true after all. Contradiction.

Formally, quite the opposite, there have been many difficulties in deriving *falsum* from Yablo's list, casting doubts on its paradoxical character.

What does it mean for an expression or a class of expressions to be paradoxical? Usually we find two different answers to this question. The most popular is the idea that a set of expressions of a language is paradoxical only if it is possible to derive a contradiction from it, along with some other principles previously accepted. Is in this sense that Yablo claims that the list written above is paradoxical. The intuitive piece of reasoning he offers immediately after presenting the list supports this idea.

However, there is al least another traditional answer to our question in the literature. A second sense in which we can say a sentence or class of sentences is paradoxical is only if it is impossible to assign stable truth values to them, again taking into account some other previously admitted principles.

Throughout this paper we will favor the first sense of 'paradoxicality' over the other, though not discarding this last sense. Hence, two questions are to be answered. Is the list of sentence that constitutes the Yablo Paradox paradoxical in our first sense? Is it in our second sense?[1]

2 The Yablo Paradox in first order languages

Sentences in Yablo's original list refer to each other by means of the position they have been assigned in the sequence: a natural number. Most of the informal reasoning makes use of some properties of the ordering of the natural numbers. Since it is possible to express every sentence of the list in a first order language, most formalizations have been developed in the language of Peano first order Arithmetic.

2.1 Definitions and notational conventions

Let $\mathcal{L}$ be the language of Peano first order Arithmetic with symbols for all primitive recursive functions and relations. If we expand $\mathcal{L}$ by adding a monadic predicate symbol T for representing the truth predicate, we obtain the language $\mathcal{L}_T$.

Let $\mathcal{N}$ be $\mathcal{L}$'s intended model, and let ω be its domain. Let $\mathcal{PA}$ be the usual axiomatization of Peano Arithmetic formulated within $\mathcal{L}$, containing the Induction Schema and equations defining primitive recursive functions and relations in a natural way.[2] Let

[1] The literature regarding the Yablo Paradox deals with three main topics: the paradoxical character of the list, which specially concerns us, the circular character of the list, which is the most popular issue but we will only devote a few lines to it, and whether the paradox is liar-like or not. This matter will be avoided here. For information and a discussion over this topic, see for instance Yablo [17] and Hardy [7].

[2] This is, in the intended interpretation.

$\mathcal{PA}_T$ be the theory in $\mathcal{L}_T$ that obtains by adding to $\mathcal{PA}$ all instances of the Induction Schema generated by $\mathcal{L}_T$ formulas containing the predicate symbol T.

If φ is a formula of $\mathcal{L}$ or any expansion of it, $\ulcorner\varphi\urcorner$ is the numeral of the Gödel number of φ. The substitution function, when applied to (the Gödel number of) a formula x, (the Gödel number of) a term t and (the Gödel number of) a variable v, gives (the Gödel number of) the formula that results from replacing the free variable (whose Gödel number is) v in the formula (whose Gödel number is) x by the term (whose Gödel number is) t. This ternary function is primitive recursive. Therefore, it will be represented in $\mathcal{PA}$ (and in any extension of it), say, by the ternary function expression $x(t/v)$.[3] The function that maps each number to its numeral is also primitive recursive and will be represented within $\mathcal{PA}$ (and in any extension of it) by $\dot{x}$. Following Feferman's 'dot' notation, if φ is a formula with exactly one free variable v, we write $\forall x T\ulcorner\varphi(\dot{x})\urcorner$ as short for $\forall x T\ulcorner\varphi\urcorner(\dot{x}/v)$ to bind variables that occur within closed terms (numerals of Gödel numbers of formulas, for instance). Then, in a standard interpretation, $\forall x T\ulcorner\varphi(\dot{x})\urcorner$ states that every numerical instance of φ is true.

2.2 Formalizing the Yablo Paradox

A natural way of doing so is to follow the strategy used for the Arithmetical Liar,[4] this is, utilizing biconditionals instead of the identity symbol to allow sentences to refer to themselves or other sentences. Each expression of the formalized list will then be a member of

$$\mathcal{YA} = \{Y(\bar{n}) \leftrightarrow \forall x(x > \bar{n} \to \neg T\ulcorner Y(\dot{x})\urcorner) : n \in \omega\}^5$$

There are at least two alternatives for constructing the list in an arithmetical first order language,[6] both due to Priest [11].

Graham Priest's first proposal is to formalize it within $\mathcal{L}_T$ by means of the Diagonalization Lemma. Since $\mathcal{PA}_T$ is a finite extension of $\mathcal{PA}$, we can apply the standard generalization of the Diagonalization Lemma for formulas with exactly two free variables[7] to $\forall x(x > z \to \neg Ty(\dot{x}/z))$ and get:

$$\mathcal{PA}_T \vdash \forall z(Y(z) \leftrightarrow \forall x(x > z \to \neg T\ulcorner Y(\dot{x})\urcorner))$$

This theorem of $\mathcal{PA}_T$ is what Ketland [8] calls 'The Uniform Fixed-Point Yablo Principle' (UFPYP, from now on), and it shows that the Yablo predicate is a fixed-point

[3] Notice that the predicates 'x is a formula', 'x is a term' and 'x is a variable' are primitive recursive too. Also, we are assuming that, if necessary, bound variables in x will be renamed in order to avoid unintended bindings.

[4] The Arithmetical Liar is given by the following sentence of $\mathcal{L}_T$: $\lambda \leftrightarrow \neg T\ulcorner\lambda\urcorner$.

[5] As usual, $\bar{n}$ is the numeral of n for each $n \in \omega$.

[6] By 'arithmetical languages' I mean those extensions of $\mathcal{L}$ that do not add new terms to the language.

[7] According to this generalization of the lemma, if ψ is a formula of $\mathcal{L}_T$ with exactly two free variables x and y, there is another formula φ of $\mathcal{L}_T$ with exactly one free variable y such that

$$\mathcal{PA}_T \vdash \forall y(\varphi(y) \leftrightarrow \psi(\ulcorner\varphi(y)\urcorner, y))$$

of $\forall x(x > z \to \neg Ty(\dot{x}/z))$. Yablo's list is derivable from the UFPYP just by instancing the universal quantifier with each numeral of the language.

However, this is not a good path, if we believe that the Yablo Paradox is not circular in its original presentation.[8] For, although it gives us a formal paradox, this paradox turns out to be circular. In the next few lines I will argue in favor of this idea.

Historically, paradoxes have been called intuitively 'circular' when constituted by expressions that, directly or indirectly, refer to themselves in some way or other. The sentence

(L) L is untrue.

—core of what has been called 'the Liar Paradox'—directly states something about itself. Each of the following three sentences

(L_1) L_2 is untrue.

(L_2) L_3 is untrue.

(L_3) L_1 is untrue.

on the other hand, refers to itself indirectly, by means of the other two. Therefore, we can conclude that the paradox constituted by them is circular, at least in a intuitive sense.

Semantic antinomies of this sort can usually be expressed within any extension of first order arithmetic containing a truth predicate. Circular expressions involved in them are traditionally—and suitably—formalized by means of what I will call '*diagonal sentences*', *i.e.*, expressions that just state, explicitly, that a sentence or predicate is a weak fixed-point of some other predicate:[9]

- If α is a sentence and ψ is a formula with exactly one free—individual—variable,

$$\alpha \leftrightarrow \psi(\ulcorner \alpha \urcorner)$$

is a '*diagonal sentence*'. The core of the formalized Liar Paradox in $\mathcal{PA}_T$ is given by the diagonal sentence

$$\lambda \leftrightarrow \neg T(\ulcorner \lambda \urcorner)$$

which is a theorem of $\mathcal{PA}_T$. It shows that λ is a weak fixed-point of $\neg T(x)$.

[8]Besides Yablo, Tennant [14] and Sorensen [13] argue in favor of this idea.

[9]See Cook [3].

- If φ is a formula with exactly n free—individual—variables $v_1, ..., v_n$ and ψ another formula with exactly $n+1$ free—individual—variables $v, v_1, ..., v_n$,

$$\forall v_1 ... \forall v_n (\varphi(v_1, ..., v_n) \leftrightarrow \psi(\ulcorner \varphi(v_1, ..., v_n) \urcorner, v_1, ..., v_n))$$

  is a '*diagonal sentence*'. According to this definition,

$$\forall z (Y(z) \leftrightarrow \forall x (x > z \rightarrow \neg T \ulcorner Y(z) \urcorner (\dot{x}/z)))$$

  is a diagonal sentence of $\mathcal{PA}_T$, for it is a theorem of this arithmetical system.[10] It explicitly states that Y is a weak fixed-point of $\forall x(x > z \rightarrow \neg Ty(\dot{x}/z))$.

- Nothing else is a *diagonal sentence*.

Diagonal sentences provide us means for expressing circular statements within arithmetical languages. Any time the core of a paradox formalized in such a language contains a diagonal sentence, or entails one by logic alone, it seems correct to say that the formal paradox is circular.[11] Therefore, any time some expressions belonging to the core of a paradox obtain by a single application of the Diagonalization Lemma, it seems right to assert that the paradox is circular.

Priest [11] has argued in favor of the circularity of the paradox obtained through the Diagonalization Lemma. But for the wrong reasons. He claims that '[...] the paradox concerns a predicate, $\dot{s}$ $[Y]$,[12] of the form $k > x, \neg S(k, \dot{s})$ $[\forall x(x > z \rightarrow \neg T \ulcorner Y(z) \urcorner (\dot{x}/z))]$; and the fact that $\dot{s} =$ '$k > x, \neg S(k, \dot{s})$' $[\forall z(Y(z) \leftrightarrow \forall x(x > z \rightarrow \neg T \ulcorner Y(z) \urcorner (\dot{x}/z)))]$ shows that we have a fixed-point, $\dot{s}$ $[Y]$, here, of exactly the same self-referential kind as in the Liar Paradox. In a nutshell, $\dot{s}$ $[Y]$ is the predicate 'no number greater than x satisfies this predicate'. The circularity is now manifest.' [11, p. 238]

Then, according to Priest, the paradox is circular for the reason that Y is a weak fixed-point of some predicate. But, as Cook [3] notices, every predicate and every sentence of (any extension of) $\mathcal{L}$ is a weak fixed-point of some predicate within (an extension of) PA (in that language). It seems reasonable to think that not every arithmetical statement is circular. However, if we adopt Priest's *criterion*, *i.e.*, if we agree that an expression is circular any time it contains predicates that are fixed-points of some other predicates, we commit ourselves with the circularity of every expression. Hence, including a predicate that is a fixed-point of some other is not enough for a statement to be circular—as Cook and Priest claim. What clearly is, as already stated, is being a diagonal sentence; the Yablo Paradox formalized via Diagonalization Lemma is circular because its formulation involves a diagonal sentence, the UFPYP.

[10] By 'arithmetical system' or 'arithmetical theory' I mean any extension of $\mathcal{PA}$, first or second order.

[11] Note that this condition, if adequate, is only a sufficient but not a necessary one. It is easy to see that a natural formalization of sentences (L_1)-(L_3) presented above within $\mathcal{L}_T$ logically entails a diagonal sentence and is, thus, straightforwardly circular.

[12] Priest utilizes a slightly different notation that makes no difference regarding results. Formulas between brackets are the translation of his notation to ours and, of course, do not appear in the original.

In fact, as Priest [11] and Ketland [9] show, adding to $\mathcal{PA}_T$ principles governing the truth predicate that are strong enough to prove what Ketland calls the 'Uniform Yablo Disquotation Principle' (UYDP, from now on):

$$\forall x(T\ulcorner Y(\dot{x})\urcorner \leftrightarrow Y(x))$$

allows us to derive an inconsistency from the list, but only making use of the UFPYP, a diagonal sentence. This principle is not merely arithmetical, nor truth-theoretical. Neither is just both. It is a sentence that states something about the behavior of the Yablo predicate, something that the biconditionals in the list by themselves do not entail (as we will see immediately) and, therefore, belongs to the core of the formal paradox. Since it is a diagonal sentence, we must regard the paradox as circular.

Actually, the UFPYP entails all Yablo biconditionals but this does not hold the other way around. There is an alternative way for obtaining the list, also suggested by Priest [11] and developed by Ketland [8], that blocks the UFPYP. Let $\mathcal{L}_{TY}$ be the language that results from adding to $\mathcal{L}_T$ a new monadic predicate symbol Y; and let $\mathcal{PA}_{TY}$ be the theory in this language that obtains by incorporating to $\mathcal{PA}_T$ all members of the set $\mathcal{YA}$ and all instances of the Induction Schema generated by formulas of $\mathcal{L}_{TY}$ containing Y. The UFPYP is no longer a theorem of $\mathcal{PA}_T$, nor of $\mathcal{PA}_{TY}$.[13] Of course, Y will be a fixed-point of some binary predicate ψ, but different from $\forall x(x > z \rightarrow \neg Ty(\dot{x}/z))$; and the diagonal sentence stating that will no longer belong to the core of the paradox: only the Yablo biconditionals. Since none of these are diagonal sentences themselves, there is no reason to think that Yablo's list is circular.

Thus, adding to $\mathcal{PA}_T$ the list of Yablo biconditionals by means of a new monadic predicate Y is a better way of guaranteeing the existence of the list, for two reasons. In the first place, it does not allow principles governing Y other that the Yablo biconditionals in the core of the paradox, just like Yablo's [16] original presentation does not allow anything to constitute the paradox but the sentences in the list. Our purpose is to analyze the possibility of getting a contradiction from the list, along with reasonable arithmetical and truth-theoretical principles, not along with the UFPYP or any other axioms governing Y. Secondly, this gives us no motive to regard the formalized list as circular, just like in the semi-formalized case. Our special interest in the Yablo Paradox is, as stated at the beginning, its non-circular character. A circular formalization of the non-circular original version is worthless, and does not deserve to be regarded as a real formalized version of the Yablo Paradox. Thus, I will leave it behind.

Nonetheless, this is not a promising path either. For as Ketland [8] notices, no matter which truth principles we add to $\mathcal{PA}_{TY}$, the resulting theory will be consistent, as long as they are not inconsistent by their own within $\mathcal{PA}_{TY}$, without the list of formalized sentences that form the Yablo Paradox. As already said, it is only possible to get a contradiction from the list along with the UFPYP; but the Yablo biconditionals by themselves do not entail this principle.

[13]For Y is no longer the predicate we get by applying the Diagonalization Lemma to $\forall x(x > z \rightarrow \neg Ty(\dot{x}/z))$.

This alternative way of getting Yablo's list of sentences within an arithmetical language leading to a consistent theory is a consequence of the Compactness Theorem, as Forster [6] first notices. Let $\mathcal{YT}$[14] be the theory that obtains by adding to $\mathcal{PA}_{TY}$ the UDYP. Without members of $\mathcal{YA}$, the theory has a standard model; and also together with any finite subset of $\mathcal{YA}$. Hence, by Compactness, the whole $\mathcal{YT}$ is satisfiable. The intuitive argument for a contradiction from de semi-formalized sentences relies on the infiniteness of the list and, since first order languages are compact, it cannot be mirrored by them. This pushes us, not to the circular formalization, but to considering different—more powerful —logical systems, where the Compactness Theorems fails. This is the subject of section 3.

The list of Yablo's sentences in a first order language is not paradoxical in our first sense of the term; it is impossible to derive *falsum* from the list. Neither is the list paradoxical in the second sense mentioned above. For, by Compactness, as already seen, there must be a model for $\mathcal{YT}$. The existence of this model entails the possibility of assigning stable truth values to every sentence of the form $Y(\bar{n})$ with $n \in \omega$, for models are complete. We must conclude, then, that Yablo's list is not paradoxical formulated within first order languages.

2.3 Circularity or consistency and ω-inconsistency

So far, we have two alternative formalizations of the Yablo Paradox. The first one, though inconsistent, seems to be clearly circular and should be discarded. The second one, as far as we know, non-circular, turns out to be consistent. Neither way gets us a non-circular antinomy.

But there is more to be said about the second formalization of Yablo's list. As Hardy [7] notices and Ketland [9] proves, $\mathcal{YT}$ is an ω-inconsistent theory. For we can prove in that system every numerical instance of $\neg Y(x)$ and also the formula $\exists x Y(x)$. Hence, since the UDYP seems to be a reasonable principle for handling the truth predicate for the sentences in $\mathcal{YA}$—Hardy and Ketland conclude—Yablo's list is not a paradox but an ω-paradox.[15]

The Compactness Theorem entails the consistency of the list even along with reasonable arithmetical and truth-theoretical principles. The consistency but ω-inconsistency of the list is guaranteed by the existence of non-intended models of $\mathcal{YT}$, that allow every numerical instance of $Y(x)$ to be true and, yet, its universal closure to be false. Second order languages with full semantics are not compact, nor allow non-intended models. Thus, the possibility of getting a contradiction from the list within a second order language is still an open question.

[14] For 'Yablo Theory', naturally.

[15] Reasonable as the UDYP may seem, Leitgeb [10, p. 71] argues that an ω-inconsistent truth theory '[...] may no longer be interpreted as speaking (only) about natural numbers and thus about the codes of sentences, but rather about nonstandard numbers, and codes of nonstandard sentences.', since it has no ω-model; and Barrio [1] concludes, thus, that T cannot be a truth predicate for the language after all. This oddity is precisely what means to be an ω-paradox. Though Yablo's first order version does not entail inconsistency, it causes what Leitgeb [10, p. 71] calls '[...] a drastic deviation from the intended ontology of the theory.'

3 Ω-inconsistent second order theories

Before examining a second order version of the Yablo Paradox, since its first order version turned out to be consistent though ω-inconsistent, it will be useful to previously consider a larger *phenomenon*: first order consistent ω-inconsistent theories's second order rewritings.

3.1 Definitions and notational conventions

Let $\mathcal{L}^2$ be a second order language containing the same non-logical symbols as $\mathcal{L}$. Let $\mathcal{N}^2$ be the intended model of $\mathcal{L}^2$. Again, ω is the range of first order quantifiers in $\mathcal{N}^2$. Also, let $\mathcal{PA}^2$ be the second order theory that results from replacing all instances of the Induction Schema of $\mathcal{PA}$ by the second order Induction Axiom.[16]

Let L^1 be a first order extension of $\mathcal{L}$ and L^2 be the second order language with the same non-logical vocabulary as L^1. If A^1 is a first order extension of $\mathcal{PA}$ expressed in L^1, I will call 'A^1's *second order rewriting*' the second order theory A^2 that results from replacing homogeneously the schematic n-ary letters that appear in A^1's axioms schemata with n-ary set variables and then bounding them universally from outside. Every other axiom of A^1 remains intact. Manifestly, $\mathcal{PA}^2$ is $\mathcal{PA}$'s second order rewriting.

3.2 Looking for a general result

It is not clear whether second order rewritings of first order consistent but ω-inconsistent theories are consistent or not. Certainly, they lack of models as a result of Categoricity.[17]

Let A^2 be an extension of $\mathcal{PA}^2$ in a second order language L^2. Then:

Theorem 3.1 *If there is a model $\mathcal{M}$ such that $\mathcal{M} \vDash A^2$ and, for some formula $\varphi(x)$ of L^2, $\mathcal{M} \vDash \varphi(\bar{n})$ for each $n \in \omega$, then $\mathcal{M} \vDash \forall x \varphi(x)$.*

Proof *Since A^2 contains as theorems every axiom of $\mathcal{PA}^2$, if $\mathcal{M} \vDash A^2$ then $\mathcal{M} \vDash \mathcal{PA}^2$ and, by Categoricity, $\mathcal{M}$ is isomorphic to $\mathcal{N}^2$. Then, if for some formula $\varphi(x)$ of L^2 $\mathcal{M} \vDash \varphi(\bar{n})$ for each $n \in \omega$, then $\mathcal{N}^2 \vDash \varphi(\bar{n})$ for each $n \in \omega$ too. As $\mathcal{N}^2$'s first order domain is ω, $\mathcal{N}^2 \vDash \forall x \varphi(x)$ and, by isomorphism, also $\mathcal{M} \vDash \forall x \varphi(x)$.* □

Theorem 3.1 entails the soundness of the ω-rule[18] within $\mathcal{PA}^2$. This, in turn, implies the next result, being A^2 as before:

Corollary 3.2 *If A^2 is an ω-inconsistent theory, it is unsatisfiable.*

[16] For a precise definition of 'second order language', 'second order theory', second order semantics and second order arithmetic, see Shapiro [12]. I will only consider second order theories without the Axiom of Choice, since there are many concerns around its logical status and it will not play a significant role here.

[17] Of course, considering standard semantics. The first categoricity result is due to Dedekind [4].

[18] The ω-rule allows us to infer $\forall x \varphi(x)$ from the set $\{\varphi(\bar{n}) : n \in \omega\}$ for any formula φ of the language.

Proof *If A^2 is ω-inconsistent, there is a formula $\varphi(x)$ of L^2 such that $A^2 \vdash \neg\varphi(\bar{n})$ for each $n \in \omega$ and, also, $A^2 \vdash \exists x\varphi(x)$. Suppose for contradiction that there is a model $\mathcal{M}$ such that $\mathcal{M} \vDash A^2$. Then, $\mathcal{M} \vDash \exists x\varphi(x)$ and, simultaneously, $\mathcal{M} \vDash \neg\varphi(\bar{n})$ for each $n \in \omega$. But, by Theorem 3.1 this entails that $\mathcal{M} \vDash \forall x\neg\varphi(x)$, which is impossible.* □

Corollary 3.2 entails the unsatisfiability of every second order rewriting of a consistent but ω-inconsistent theory. However, this is not enough for inconsistency, since second order logic is incomplete. It may perfectly happen that a second order theory semantically implies *falsum* but, for the failure of Completeness, *falsum* is not a theorem in it. Nonetheless, this is not the general case.

Theorem 3.3 *Some second order rewritings of first order consistent but ω-inconsistent theories are inconsistent.*

Proof *Let G be $\mathcal{PA}$'s Gödel sentence:*

$$G \leftrightarrow \forall x\neg Bew(\ulcorner G \urcorner, x)^{19}$$

As it is well known, $\mathcal{PA}$ proves $\neg Bew(\ulcorner G \urcorner, \bar{n})$ for each $n \in \omega$, but it does not prove G. Consider then the first order system $\mathcal{PA} \cup \{\neg G\}$. This theory is consistent but ω-inconsistent. Its second order rewriting, $\mathcal{PA}^2 \cup \{\neg G\}$, proves G as a theorem, since $\mathcal{PA}^2$ does, and is, therefore, inconsistent simpliciter. □

If the result that Theorem 3.3 provides us could be generalized to every first order consistent but ω-inconsistent system, Yablo's list of sentences would constitute a real paradox in our first and privileged sense of the term. Could this be our general result?

Unfortunately, no, as the following theorem shows. Let A^1 be any first order extension of $\mathcal{PA}$ in L^1 and let $\varphi(x)$ be a formula of this language. Also, let A^2 be A^1's second order rewriting, formulated in L^2, and let $E = \{\varphi(\bar{n}) : n \in \omega\}$.

Theorem 3.4 *If for any finite set $F \subseteq E$ there is an extension $\mathcal{N}_F$ of $\mathcal{N}$ such that $\mathcal{N}_F \vDash A^1 \cup F$, then $A^2 \cup E \nvdash \bot$.*

Proof *I will show that, for every finite $F \subseteq E$, $A^2 \cup F$ has a corresponding model $\mathcal{N}^2_F$. Given such an F, we extend $\mathcal{N}^2$ to a model $\mathcal{N}^2_F$ of L^2 in the following way:*[20]

- *if c is an individual constant of L^2, $c^{\mathcal{N}^2_F} = c^{\mathcal{N}_F}$;*
- *if f is an n-ary function symbol of L^2, $f^{\mathcal{N}^2_F} = f^{\mathcal{N}_F}$;*
- *if P is an n-ary predicate symbol of L^2, $P^{\mathcal{N}^2_F} = P^{\mathcal{N}_F}$;*

[19] As usual, $Bew(x, y)$ represents in (any extension of) $\mathcal{PA}$ the relation 'y is the Gödel number of a proof within (that extension of) $\mathcal{PA}$ of the formula whose Gödel number is x'.

[20] Regarding Model Theory, I will follow notation from Boolos *et al.* [2]. This detailed extension is entirely possible, because $\mathcal{N}_F$ has ω as its domain, since it is an extension of $\mathcal{N}$ by hypothesis.

Now I will show that $\mathcal{N}_F^2 \vDash A_2 \cup F$. Firstly, since $\mathcal{N}_F^2$'s first order domain is ω and it interprets exactly in the same way as $\mathcal{N}_F$ every non-logical symbol of L^1, for every L^1 formula α such that $\mathcal{N}_F \vDash \alpha$, $\mathcal{N}_F^2 \vDash \alpha$. Then, $\mathcal{N}_F^2 \vDash A^1 \cup F$. Secondly, both the Induction Axiom and the Comprehension Schema are true in $\mathcal{N}_F^2$. This is trivial, since they were true in $\mathcal{N}_2$. The only thing that has changed by extending this model to $\mathcal{N}_F^2$ regarding those axioms is that the later model may interpret some new n-ary predicate symbols by assigning them sets of n-tuples of members of ω that were already part of the range of the set variables. It does not matter if they used to have no name. Consequently, $\mathcal{N}_F^2 \vDash A^2 \cup F$.

Now assume for contradiction that $A^2 \cup E \vdash \bot$. By the Finiteness Theorem we know that only a finite number of members of E can be utilized in the proof. Let F be the finite subset of such members of E. Then, $A^2 \cup F \vdash \bot$, which is absurd, for we have shown that every $A^2 \cup F$ has a model. □

Theorem 3.4 implies that, even $A^1 \cup E$ being an ω-inconsistent system, if for every finite $F \subseteq E$ $A^1 \cup F$ is ω-consistent, then $A^2 \cup E$, $A^1 \cup E$'s second order rewriting, is consistent. Since, as we will see in the next section, such theories exist, we must conclude that not every ω-inconsistent first order system has an inconsistent second order rewriting.

Consequently, the search for a general result is hopeless. While some second order rewritings of ω-inconsistent first order theories turn out to be inconsistent, others do not. However, Theorem 3.4 gives us a general rule, for it shows that, any time the ω-inconsistent character of a first order theory is a product of adding infinitely many sentences—it is a necessary and sufficient condition for ω-inconsistency—, its second order rewriting is then consistent.

Theorem 3.3's result, on the contrary, cannot be generalized, *i.e.*, it is not the case that any time we get ω-inconsistency by adding only a finite number of formulas to an arithmetical system, the second order rewriting of the resulting theory in inconsistent. If it were, $\mathcal{PA}^2$ would prove every true Π_1^0 statement, which it does not.[21]

4 The Yablo Paradox in second order languages

As in the first order case, there are two ways of guaranteeing the existence of the list: one via the Diagonalization Lemma, which gives us a circular paradox and is, therefore, useless; and another one, that introduces Y as a new primitive in the language. I will stick to this second path.

Our concern here is firstly with the second order rewriting of $\mathcal{YT}$. As this system is consistent but ω-inconsistent, some of the results of the previous section will become handy.

[21] I am very grateful with Roy T. Cook and, indirectly, with Stewart Shapiro for pointing this out to me.

4.1 Definitions and notational conventions

Let $\mathcal{YT}^2$ be $\mathcal{YT}$'s second order rewriting and let $\mathcal{L}^2_{TY}$ be the language in which $\mathcal{YT}^2$ is formulated.

Regarding axioms, $\mathcal{YT}^2$ differs from $\mathcal{YT}$ just in the induction case. The former replaces every instance of $\mathcal{YT}$'s Induction Schema with a single statement, the Induction Axiom. This arithmetical principle, along with the Comprehension Schema, allows us to derive every first order instance of the Induction Schema. Thus, all $\mathcal{YT}$ axioms are theorems of $\mathcal{YT}^2$ and, consequently, as $\mathcal{YT}$ is ω-inconsistent, $\mathcal{YT}^2$ is so too. Since the Compactness Theorem does not hold for $\mathcal{YT}^2$ and, by the Categoricity Theorem, there are no non-intended models to guarantee the consistency of $\mathcal{YT}^2$, we may feel inclined to believe that it is inconsistent *simpliciter*, like in the case of $\mathcal{PA} \cup \{\neg G\}$'s second order rewriting.

4.2 Bad and good news

Unfortunately, like in first order, second order calculus is not powerful enough to mirror the original informal reasoning and get us a contradiction from $\mathcal{YT}^2$. $\mathcal{YT}^2$'s case is not analogous to $\mathcal{PA}^2 \cup \{\neg G\}$'s. On the contrary, it fits Theorem 3.4's hypothesis, as the next corollary establishes.

Corollary 4.1 *$\mathcal{YT}^2$ is a consistent system.*

Proof *Let $\mathcal{PA} \cup \{UDYP\}$ (with all instances of the Induction Schema generated by formulas of $\mathcal{L}_{TY}$) take the place of A_1 in Theorem 3.4 and let $\mathcal{YA}$ take the place of E. Given any finite $F \subseteq \mathcal{YA}$ we extend $\mathcal{N}$ to $\mathcal{N}_F$ in the following way:*

- *If $F = \phi$, let $Y^{\mathcal{N}_F} = T^{\mathcal{N}_F} = \phi$.*
- *If $F \neq \phi$, since it is finite, there is an $m \in \omega$ such that*

  $$Y(\bar{m}) \leftrightarrow \forall x(x > \bar{m} \rightarrow \neg T \ulcorner Y(\dot{x}) \urcorner) \in F$$

  and, for all $n \in \omega$ such that $n > m$,

  $$Y(\bar{n}) \leftrightarrow \forall x(x > \bar{n} \rightarrow \neg T \ulcorner Y(\dot{x}) \urcorner) \notin F$$

  Thus, let $Y^{\mathcal{N}_F} = \{m\}$ and $T^{\mathcal{N}_F} = \{$the Gödel number of $Y(\bar{m})\}$.

Along next lines I will prove that $\mathcal{N}_F \vDash \mathcal{PA} \cup \{UDYP\} \cup F$ for any finite $F \subseteq \mathcal{YA}$. First, every $\mathcal{PA}$ axiom that does not contain T or Y is true in $\mathcal{N}_F$, for it is in $\mathcal{N}$. Secondly:

- *If $F = \phi$, $\mathcal{N}_F \vDash \forall x(Y(x) \leftrightarrow x \neq x) \wedge \forall x(T(x) \leftrightarrow x \neq x)$. Thus, every instance of the Induction Schema containing T or Y is true, since the ones containing $x \neq x$ instead are. Also, the UDYP comes out true in $\mathcal{N}_{\mathcal{F}}$, as $\forall x(T \ulcorner Y(\dot{x}) \urcorner \leftrightarrow x \neq x)$ does, because $\mathcal{N}_F \vDash \forall x \neg T(x)$.*

- *If $F \neq \phi$, $\mathcal{N}_F \models \forall x(Y(x) \leftrightarrow x = \bar{m}) \wedge \forall x(T(x) \leftrightarrow x = \ulcorner Y(\dot{x})\urcorner)$. Again, every instance of the Induction Schema containing T or Y is true, for the ones containing $x = \bar{m}$ and $x = \ulcorner Y(\dot{x})\urcorner$ instead, respectively, are.*[22] *$\mathcal{N}_{\mathcal{F}} \models$ UDYP, as $\mathcal{N}_F \models T\ulcorner Y(\bar{m})\urcorner \wedge Y(\bar{m})$ and $\mathcal{N}_F \models \neg T\ulcorner Y(\bar{n})\urcorner \wedge \neg Y(\bar{n})$ for all $n \in \omega$ such that $n \neq m$. Finally, $\mathcal{N}_F \models F$. Let $Y(\bar{n}) \leftrightarrow \forall x(x > \bar{n} \rightarrow \neg T\ulcorner Y(\dot{x})\urcorner) \in F$. If $n = m$, both $\mathcal{N}_F \models Y(\bar{n})$ and $\mathcal{N}_F \models \forall x(x > \bar{n} \rightarrow \neg T\ulcorner Y(\dot{x})\urcorner)$, since the only member of $Y^{\mathcal{N}_F}$ is m. If, instead, $n \neq m$, then both $\mathcal{N}_F \models \neg Y(\bar{n})$ and $\mathcal{N}_F \models \neg\forall x(x > \bar{n} \rightarrow \neg T\ulcorner Y(\dot{x})\urcorner)$, for the same reasons as in the previous case.*

Hence, by Theorem 3.4, $\mathcal{YT}^2 \nvdash \bot$. □

The strong metatheorem we used to prove Theorem 3.4 and, thus, Corollary 4.1 is the Finiteness result for second order calculus. According to it, if we cannot derive a contradiction from any finite subset of an infinite set, we are not deriving it from the whole set either. This is the Yablo Paradox's case: any finite subset of Yablo biconditionals has a model and, thus, is not inconsistent. Hence, neither is the entire list.

The Finiteness Theorem allows us to extend what we established in Corollary 4.1 for $\mathcal{YT}^2$ to any other theory that considers different truth principles for handling Yablo biconditionals, even strictly second order ones. Let A^2 be any arithmetical second order system in $\mathcal{L}^2_{TY}$, and let $\mathcal{TT}$[23] be any set of principles governing the truth predicate, formulated in the same language.

Corollary 4.2 *If, for every finite $F \subseteq \mathcal{YA}$, $A^2 \cup \mathcal{TT} \cup F \nvdash \bot$, then $A^2 \cup \mathcal{TT} \cup \mathcal{YA} \nvdash \bot$ either.*

Proof *Straightforward, by the Finiteness Theorem.* □

Therefore, if we manage to get *falsum* from Yablo's list within a second order language, two things may happen. Either our arithmetical and truth-theoretical principles are inconsistent by themselves—in which case we cannot know for sure if the list is to blame for the consistency, *i.e.*, if it is paradoxical or not in our first sense of the term—; or only a finite number of Yablo biconditionals is necessary for deriving a contradiction from the list—in which case we are diverting from Yablo's original semi-formalized piece of reasoning in a significant and undesirable way. No matter which truth principles we embrace, as long as they are reasonable,[24] the resulting theory will be consistent.

Nonetheless, $\mathcal{YT}^2$ is paradoxical in our second sense of the term, since it semantically entails a contradiction, as the following result shows.

Corollary 4.3 *$\mathcal{YT}^2$ is unsatisfiable.*

[22] Although Y is not a predicate of $\mathcal{L}$, $\ulcorner Y(t)\urcorner$, where t is any term of $\mathcal{L}$, is a term of $\mathcal{L}$, since it is just a numeral.

[23] For 'truth theory'.

[24] This is, inconsistent nor with the arithmetical base theory, nor with any finite subset of Yablo biconditionals.

Proof *Straightforward, by Corollary 3.2.* □

The lack of models for $\mathcal{YT}^2$ implies that, assuming $\mathcal{PA}^2$ and the UDYP, there is no way of assigning stable truth values to Yablo's sentences; this is, sentences of the form $Y(\bar{n})$, that state that $\forall x(x > \bar{n} \rightarrow \neg T\ulcorner Y(\dot{x})\urcorner)$. For, since $\mathcal{PA}^2 \cup$ UDYP has a model, as proved in Corollary 4.1, $\mathcal{YT}^2$'s unsatisfiability shows that it is not possible to regard as true every Yablo biconditional and, hence, at least one of them must be false. Then, for that biconditional, say $Y(\bar{m}) \leftrightarrow \forall x(x > \bar{m} \rightarrow \neg T\ulcorner Y(\dot{x})\urcorner)$, either $Y(\bar{m})$ receives the truth value *truth* and what it says, $\forall x(x > \bar{m} \rightarrow \neg T\ulcorner Y(\dot{x})\urcorner)$, *false*; or, conversely, $Y(\bar{m})$ gets the truth value *false* but $\forall x(x > \bar{m} \rightarrow \neg T\ulcorner Y(\dot{x})\urcorner)$ is assigned *true*.

5 Conclusions

Formulated in a first order language, the Yablo Paradox is not paradoxical in either of the two senses of the term mentioned in the Introduction. Neither is possible to derive a contradiction from the list along with reasonable arithmetical and truth-theoretical principles within first order calculus, nor is the formalized version of Yablo's original piece of reasoning a valid argument in first order logic.

Second order calculus does not allow us to get a contradiction from the list either, whenever we choose reasonable truth and arithmetical principles to do so. Thus, Yablo's sequence is not paradoxical in our first sense within second order logic. We may then feel tempted to regard Yablo's original argument as logically incorrect, as invalid *simpliciter*; and Yablo's list of sentences as non-paradoxical. However, as Corollary 4.3 shows, this path is forbidden, since the argument is second order semantically valid.[25] If we embrace the second order notion of logical consequence we must subscribe to the idea that second order calculus is not powerful enough for representing Yablo's argument, and nor is first order calculus.

Second order logic has an advantage over first order logic. Regarding the intuitive reasoning as semantically valid, firstly, it proves $\mathcal{YT}^2$ unsatisfiability and, thus, the impossibility of assigning stable truth values to Yablo's sentences. Consequently, it regards Yablo's sequence as paradoxical in our second sense of the term. And, secondly, it is expressive enough to mirror Yablo's original argument. On the contrary, first order logic cannot do any of that.

References

[1] Barrio, Eduardo A., 'Theories of truth without standard models and Yablo's sequences', *Studia Logica*, 96 (2010), 375–391.

[2] Boolos, G., J. Burgess, and R. Jeffrey, *Computability and Logic*, Quinta edn., Cambridge University Press, Cambridge, 2007.

[25] If the reader has followed this piece of work to this point, she cannot dismiss the second order notion of logical consequence, for examining the paradoxicality of Yablo's list in second order systems has been established as our main goal at the beginning.

[3] Cook, Roy T., 'There are non-circular paradoxes (but Yablo's Isn't One of Them!)', *The Monist*, 89 (2006), 1, 118–149.

[4] Dedekind, Richard, 'Was sind und was sollen die zahlen?', in William B. Ewald, (ed.), *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, Oxford University Press, 1996, pp. 787–832.

[5] Ewald, William B., (ed.) *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, Oxford University Press, 1996.

[6] Forster, Thomas, 'The significance of Yablo's Paradox without self-reference', , 1996.

[7] Hardy, James, 'Is Yablo's Paradox liar-like?', *Analysis*, 55 (1995), 3, 197–198.

[8] Ketland, Jeffrey, 'Bueno and Colyvan on Yablo's Paradox', *Analysis*, 64 (2004), 165–172.

[9] Ketland, Jeffrey, 'Yablo's Paradox and ω-inconsistency', *Synthese*, 145 (2005), 295–307.

[10] Leitgeb, Hannes, 'Theories of truth which have no standard models', *Studia Logica*, 68 (2001), 69–87.

[11] Priest, Graham, 'Yablo's Paradox', *Analysis*, 57 (1997), 236–242.

[12] Shapiro, Stewart, *Foundations without Foundationalism: A Case for Second-Order Logic*, Oxford University Press, Nueva York, 1991.

[13] Sorensen, Roy A., 'Yablo's Paradox and kindred infinite liars', *Mind*, 107 (1998), 137–155.

[14] Tennant, Neil, 'On paradox without self-reference', *Analysis*, 55 (1995), 199–207.

[15] Yablo, Stephen, 'Truth and reflexion', *Journal of Philosphical Logic*, 14 (1985), 297–349.

[16] Yablo, Stephen, 'Paradox without self-reference', *Analysis*, 53 (1993), 251–252.

[17] Yablo, Stephen, 'Circularity and paradox', in Bolander, Hendricks, and Pedersen, (eds.), *Self-Reference*, CSLI Publications, Stanford, 2004, pp. 139–157.