Lavinia María Picollo

Yablo's Paradox in Second-Order Languages: Consistency and Unsatisfiability

Abstract. Stephen Yablo [23, 24] introduces a new informal paradox, constituted by an infinite list of semi-formalized sentences. 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 sequence, 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 is said to be paradoxical, since second-order languages are not compact, I study 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.

Keywords: Paradoxicality, Consistency, Ω-Inconsistency, Second-order languages, Unsatisfiability, Finiteness.

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). The first and most popular one was Yablo's Paradox, introduced by Stephen Yablo [23, 24]. 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. Yablo's 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.

$\ldots$

Presented by **Richmond Thomason**; *Received* October 23, 2011

Studia Logica
DOI: 10.1007/s11225-012-9399-6

Published online: 12 October 2012

Informally, it is possible to establish the paradoxicality of the list by getting a contradiction from it or showing that it is impossible to assign truth-values to its sentences in a consistent way. Suppose for *reductio* that some Y_n is true. Since it says that the following are not, they are not. Thus, 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. Hence, Y_{n+1} must be true after all, which is absurd because we have just said it was not. Thus, every Y_n in the sequence is false. In particular, all sentences below Y_0 are false and, therefore, Y_0 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 paradoxicality.[1]

What does it mean for an expression or class of expressions to be paradoxical? Usually we find two different answers to this question implicit in the literature. The most popular is the idea that a set of expressions is paradoxical only if it allows us to derive a contradiction from some intuitive and previously accepted principles.[2] In formal languages, this means that the theory entailing such principles is inconsistent, given the existence of such set of expressions. Yablo's piece of reasoning could be described as an attempt to obtain a contradiction from some arithmetical and truth-theoretical principles, assuming the existence of the sequence of sentences he introduces.

Nonetheless, there is at least another traditional answer to our question. A second sense in which we can say a class of expressions is paradoxical is only if it is impossible to assign (classic) truth-values consistently to its members, i.e., if the assumption that some of them are true makes them false and vice versa, again taking into account some intuitive and previously admitted principles.[3] In formal languages, this implies that such principles

[1] The literature regarding Yablo's Paradox deals with three main topics: the paradoxical character of the list, which specially concerns us, the circularity 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 [25] and James Hardy [12].

[2] Paul Benacerraf and Crispin Wright [3], Jeffrey Ketland [13,14], Graham Priest [18] and Neil Tennant [22], for instance, embrace this notion of paradoxicality. Sometimes this requirement is replaced with a weaker one: a (merely) false statement must be syntactically entailed. Curry's Paradox proving the existence of Santa Claus is one clear example. However, in most cases both requirements turn out to be equivalent: of course, the former implies the later; and generally the derivation of a falsity contradicts some of the previously accepted principles.

[3] For instance, Nuel Belnap and Anil Gupta [2], Roy T. Cook [5], Hartry Field [9], Ketland [14] and Saul Kripke [15] use the term 'paradoxical' in this sense.

are unsatisfiable, assuming the existence of such set of expressions. Yablo's original argument could also be described as an attempt to show the impossibility of assigning consistent truth-values to sentences in the sequence.

For languages in which the Completeness Theorem holds, e.g. first-order languages, these two senses are equivalent. However, this is not generally the case, since many formal languages, such as higher-order ones, are not complete. Throughout this paper we will favor the first sense of 'paradoxical' over the other one, though not discarding this last sense. Hence, two questions are to be answered: Is Yablo's list paradoxical in our first sense? Is it in our second sense?

2. Yablo's Paradox in First-Order Languages

Sentences in Yablo's original sequence refer to each other by means of the position they have been assigned in it: a natural number. Most of the informal reasoning makes use of some properties of the ordering of the natural numbers, as well as principles for truth. Since it is possible to express every sentence of the list in a first-order language, first formalizations have been developed in first-order arithmetical languages[4] containing a monadic predicate symbol T for truth.

2.1. Definitions and Notational Conventions

Let $\mathcal{L}$ be the language of first-order Peano Arithmetic with symbols for all primitive recursive functions and relations. If we extend $\mathcal{L}$ with T, we obtain the language $\mathcal{L}_T$.

Let $\mathbb{N}$ be $\mathcal{L}$'s intended model, and let ω, the set of natural numbers, be its domain.

Let $\mathcal{PA}$ be the usual axiomatization of first-order Peano Arithmetic formulated within $\mathcal{L}$, containing the Induction Schema and equations defining primitive recursive functions and relations in a natural way.[5] Let $\mathcal{PA}_T$ be the theory in $\mathcal{L}_T$ that obtains by adding to $\mathcal{PA}$ all instances of the Induction Schema generated by $\mathcal{L}_T$-formulae containing T.

If φ is a formula of an arithmetical language, $\ulcorner\varphi\urcorner$ is the numeral of the Gödel number of φ. The substitution function, when applied to the Gödel

[4] By '*arithmetical language*' I understand any first-order or second-order extension of the language of first-order Peano Arithmetic that does not allow new terms.

[5] Functions and relations defined in a natural way allow us to prove basic syntactic facts that otherwise would not be provable. For more precision, see Halbach [11, pp. 33–35].

numbers x of a formula, t of a term, and v of a variable, 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 arithmetical theories,[6] say, by $x(t/v)$.[7] The function that maps each number to its numeral is also primitive recursive and will be represented within arithmetical systems 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 (e.g. numerals of Gödel numbers of formulae). Thus, in a standard interpretation of the non-logical vocabulary, $\forall x T\ulcorner\varphi(\dot{x})\urcorner$ states that every numerical instance of φ is true.

2.2. Formalizing Yablo's Paradox

The usual way of doing so within arithmetical languages is to follow the strategy used for the Arithmetic Liar and the Gödel sentence,[8] that is, utilizing biconditionals instead of identity statements to allow formulae to refer to themselves or other formulae. A predicate $Y(x)$ is needed, since biconditionals hold only between formulae, not terms such as codes of expressions. $Y(x)$ may be called '*Yablo's predicate*':

$$\mathcal{YA} = \{Y(\bar{n}) \leftrightarrow \forall x(x > \bar{n} \rightarrow \neg T\ulcorner Y(\dot{x})\urcorner) : n \in \omega\}^{9}$$

For each $n \in \omega$, Yablo's original sentences Y_n are formalized by $Y(\bar{n})$, and biconditionals in $\mathcal{YA}$ set their reference, for they establish an equivalence between each of them and the claim that, for all $m > n$, $Y(\bar{m})$ is untrue.

There are at least two alternative ways of guaranteeing the existence of the list—i.e., of allowing biconditionals in $\mathcal{YA}$ to hold—in a first-order arithmetical system, both due to Priest [19].

Priest's first proposal is to formulate it in $\mathcal{PA}_T$ by means of the Diagonalization Lemma. Since this system is a recursive extension of $\mathcal{PA}$, we can

[6] By '*arithmetical system*' or '*arithmetical theory*' I understand any first-order or second-order recursive extension of $\mathcal{PA}$.

[7] 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.

[8] The Arithmetic Liar, λ, and the Gödel sentence, G, are given by '$\lambda \leftrightarrow \neg T\ulcorner\lambda\urcorner$' and '$G \leftrightarrow \forall x \neg Bew(\ulcorner G\urcorner, x)$', respectively. As usual, in an arithmetical theory $Bew(x, y)$ represents the relation 'y is the Gödel number of a proof in the system of the formula whose Gödel number is x'.

[9] For each $n \in \omega$, $\bar{n}$ is the numeral of n.

apply a standard generalization of the Diagonalization Lemma for formulae with exactly two free variables[10] to $\forall x(x > z \rightarrow \neg Ty(\dot{x}/z))$ and get:

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

This theorem of $\mathcal{PA}_T$ is what Ketland [13] calls the '*Uniform Fixed-Point Yablo Principle*' (UFPYP, from now on), and it states that Yablo's predicate is a weak fixed point of $\forall x(x > z \rightarrow \neg Ty(\dot{x}/z))$.[11] 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 Yablo's Paradox is not circular in its original presentation.[12] 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. Therefore, we can conclude that the paradox constituted by it is circular, at least in an intuitive sense. Semantic antinomies of this sort can usually be expressed within arithmetical theories containing a truth predicate; circular expressions involved in them are traditionally formalized by means of what I will call '*diagonal sentences*', i.e., expressions that just state, explicitly, that a formula is a weak fixed point of some predicate:

- If α is a sentence and $\psi(v)$ is a formula with exactly one free (individual) variable v,

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

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

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

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

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

[11] For a thorough discussion on fixed points, see Cook [6].

[12] Besides Yablo, Tennant [22] and Roy Sorensen [21] argue in favor of this idea.

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

- If $\varphi(v_1, \ldots, v_n)$ is a formula with exactly n free (individual) variables $v_1, \ldots, v_n$ and $\psi(v, v_1, \ldots, v_n)$ another formula with exactly $n+1$ free (individual) variables $v, v_1, \ldots, v_n$,

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

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. 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 it seems correct to say that the formal paradox is circular.[13]

Priest [19] 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],*[14] *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 $\dot{s} =$ '$k > x, \neg S(k, \dot{s})$' [$Y(z) \leftrightarrow \forall x(x > z \rightarrow \neg T \ulcorner Y(z) \urcorner (\dot{x}/z))$] 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.' [19, p. 238]*

[13] Note that this condition, if adequate, is only a sufficient but not a necessary one. As Hannes Leitgeb [17] has shown by reflecting on the debate on Yablo's Paradox and its alleged (non-)circularity, giving an adequate notion of circularity for first-order languages is a difficult task and any non-sophisticated explanation of the term is expected to fail. For these reasons, I will not go any deeper into this matter here.

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

Thus, according to Priest, the paradox is circular for the reason that Y is a weak fixed point of some predicate. But, as Cook [6] notices, every formula of an arithmetical language is a weak fixed point of some predicate within an arithmetical theory. It seems reasonable to think that not every such formula 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 should not be enough for circularity—as Cook and Priest claim. What clearly is, as already stated, is being a diagonal sentence; Yablo's Paradox formalized via the Diagonalization Lemma is circular because its formulation involves the UFPYP.

In fact, as Priest [19] and Ketland [14] 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 *falsum*, but only making use of the UFPYP, a diagonal sentence. This principle is not merely arithmetical, merely truth-theoretical, or merely arithmetical and truth-theoretical. It is a sentence that states something about the behavior of Yablo's predicate, something that the biconditionals in $\mathcal{YA}$ 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 to obtain the list, also suggested by Priest [19] and developed by Ketland [13], 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 formulated in this language that obtains by incorporating to $\mathcal{PA}_T$ all members of $\mathcal{YA}$ and all instances of the Induction Schema generated by $\mathcal{L}_{TY}$-formulae containing Y. The UFPYP is no longer a theorem of $\mathcal{PA}_T$, neither of $\mathcal{PA}_{TY}$.[15] Of course, Y will be a fixed point of some binary predicate, but this predicate will not be $\forall x(x > z \rightarrow \neg Ty(\dot{x}/z))$; and the diagonal sentence stating that will no longer be among Yablo's 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 set of Yablo's biconditionals by

---

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

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 biconditionals in $\mathcal{YA}$, just like Yablo's [24] original presentation does not allow anything to constitute the paradox but sentences in the list. Our purpose is to analyze the possibility of getting a contradiction from reasonable arithmetical and truth-theoretical principles[16] (or prove them unsatisfiable) along with the list, not along with the UFPYP or any other axioms governing Y. Secondly, it gives us no motive to regard the formalized list as circular, just like in the semi-formalized case. Our special interest in Yablo's Paradox lies, as stated at the beginning, on its (intuitive) non-circularity. A circular formalization of the non-circular original version is worthless, and does not deserve to be regarded as a real formalized version of Yablo's Paradox. Thus, I will leave it behind.

Nonetheless, this is not a promising path either. For as Ketland [13] notices, no matter which truth-theoretical principles we add to $\mathcal{PA}_{TY}$, the resulting theory will be consistent, as long as they are reasonable and not inconsistent on their own with $\mathcal{PA}_{TY}$, without members of $\mathcal{YA}$. As already said, it is only possible to get a contradiction from the list along with the UFPYP; but Yablo's biconditionals by themselves do not entail this principle.

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 Thomas Forster [10] first notices. Let $\mathcal{YT}$[17] be the theory that obtains by adding to $\mathcal{PA}_{TY}$ the UYDP. Without members of $\mathcal{YA}$ the theory has a standard model; and also together with any finite subset of this set. Hence, by Compactness, the whole $\mathcal{YT}$ is satisfiable. The intuitive argument for a contradiction from the semi-formalized sentences relies on the infiniteness of the list and, since first-order languages are compact, it cannot be mirrored by them. This result pushes us, not to the circular formalization, which is not a legitimate one, 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 does not allow us to derive *falsum*. 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

[16] By 'reasonable arithmetical principles' and 'reasonable truth-theoretical principles' I understand intuitively sound ones.

[17] For 'Yablo's Theory', naturally.

model entails the possibility of assigning truth-values in a consistent way to every sentence of the form $Y(\bar{n})$ with $n \in \omega$ without making any biconditional in $\mathcal{YA}$ untrue, i.e., without denying the existence of the list. We must conclude, thus, that Yablo's list is not paradoxical when formulated in first-order languages.

2.3. Circularity or Consistency and ω-Inconsistency

So far, we have two alternative formalizations of Yablo's sequence. The first one, though inconsistent, seems to be clearly circular and should be discarded. The second one, non-circular as far as we know, turns out to be consistent and satisfiable. Neither way gives us a non-circular antinomy. But there is more to be said about the second formalization of the list. As Hardy [12] notices and Ketland [14] proves, $\mathcal{YT}$ is an ω-inconsistent theory. For we can prove in that system every numerical instance of $\neg Y(x)$ and also $\exists x Y(x)$. Hence, since the UYDP seems to be a reasonable principle for handling the truth predicate applied to Yablo's sentences—Hardy and Ketland conclude—Yablo's list is not a paradox but an ω-paradox.[18]

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 $\neg Y(x)$ to be true and, yet, its universal closure to be false. Second-order languages with full semantics are not compact and do not allow such non-intended models. Thus, the possibility of getting a paradox within a second-order language is still open.

3. Ω-Inconsistent Second-Order Theories

Before examining a second-order version of Yablo's Paradox, since its first-order version turned out to be consistent though ω-inconsistent, it will be useful to consider a larger phenomenon: consistency and satisfiability of second-order versions of first-order consistent but ω-inconsistent theories.

[18] Reasonable as the UYDP may seem, Leitgeb [16, p. 71] argues that an ω-inconsistent theory of truth '[...] 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 Eduardo Barrio [1] concludes that T cannot be a truth predicate for the language after all. This oddity is precisely what it means to be an ω-paradox. Though Yablo's first-order version does not entail inconsistency, it causes what Leitgeb [16, p. 71] calls '[...] a drastic deviation from the intended ontology of the theory.'

3.1. Definitions and Notational Conventions

Let $\mathcal{L}^2$ be a second-order language containing the same non-logical vocabulary as $\mathcal{L}$. Let $\mathbb{N}^2$ be the $\mathcal{L}^2$'s intended model. Again, ω is the range of first-order quantifiers in $\mathbb{N}^2$. 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.[19]

Let L^1 be any first-order arithmetical language and L^2 be the second-order language with the same non-logical vocabulary as L^1. If A^1 is a first-order arithmetical system formulated in L^1, I will call 'A^1's *second-order rewriting*' the second-order theory A^2 that results from homogeneously replacing the schematic n-ary letters in A^1's axiom schemata with n-ary set variables and then binding 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. Certainly, they lack of models as a result of Categoricity.[20] Let A^2 be an extension of $\mathcal{PA}^2$ formulated in L^2.

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

PROOF. Since A^2 contains as theorems every axiom of $\mathcal{PA}^2$, if $\mathfrak{M} \models A^2$ then $\mathfrak{M} \models \mathcal{PA}^2$ and, by Categoricity, $\mathfrak{M}$ is isomorphic to $\mathbb{N}^2$. Thus, if for some L^2-formula $\varphi(x)$ $\mathfrak{M} \models \varphi(\bar{n})$ for each $n \in \omega$, then $\mathbb{N}^2 \models \varphi(\bar{n})$ for each $n \in \omega$ too. As $\mathbb{N}^2$'s first-order domain is ω, $\mathbb{N}^2 \models \forall x \varphi(x)$ and, by isomorphism, $\mathfrak{M} \models \forall x \varphi(x)$. ■

Theorem 3.1 entails the soundness of the ω-rule[21] within $\mathcal{PA}^2$ and any extension of it. This, in turn, implies the next result, with A^2 as before:

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

---

[19] For a precise definition of '*second-order language*', '*second-order theory*', *second-order semantics* and *second-order arithmetic*, see Stewart Shapiro [20]. 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.

[20] Of course, considering standard semantics. The first categoricity result is due to Richard Dedekind [7].

[21] 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 an L^2-formula $\varphi(x)$ 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 *reductio* that there is a model $\mathfrak{M}$ such that $\mathfrak{M} \vDash A^2$. Thus, $\mathfrak{M} \vDash \exists x \varphi(x)$ and, simultaneously, $\mathfrak{M} \vDash \neg\varphi(\bar{n})$ for each $n \in \omega$. By Theorem 3.1, $\mathfrak{M} \vDash \forall x \neg\varphi(x)$, which is impossible. ■

Corollary 3.2 entails the unsatisfiability of every second-order rewriting of a consistent but ω-inconsistent first-order arithmetical theory. However, this is not enough for inconsistency, since second-order logic is incomplete. It may perfectly happen, at least *prima facie*, that a second-order theory semantically implies *falsum* but, because of the failure of Completeness, *falsum* is not one of its theorems. 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. As is well known, $\mathcal{PA}$ proves $\neg Bew(\ulcorner G \urcorner, \bar{n})$ for each $n \in \omega$, but does not prove G.[22] 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 would constitute a 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 arithmetical system formulated 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 expansion $\mathbb{N}_F$ of $\mathbb{N}$ such that $\mathbb{N}_F \vDash A^1 \cup F$, 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 $\mathbb{N}^2_F$. Given any such F, we expand $\mathbb{N}^2$ to a model $\mathbb{N}^2_F$ of L^2 in the following way:[23]

- if c is an individual constant of L^2, $c^{\mathbb{N}^2_F} = c^{\mathbb{N}_F}$;

[22]Naturally, $\mathcal{PA} \nvdash G$ as long as it is a consistent theory. As usual, I will be working under this assumption through the whole paper. Moreover, I will assume that both $\mathcal{PA}$ and $\mathcal{PA}^2$ are arithmetically sound, i.e., that $\mathbb{N} \vDash \mathcal{PA}$ and $\mathbb{N}^2 \vDash \mathcal{PA}^2$.

[23]This detailed expansion is entirely possible, because $\mathbb{N}_F$ has ω as its domain, since it is an expansion of $\mathbb{N}$ by hypothesis.

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

Now I will show that $\mathbb{N}^2_F \models A^2 \cup F$. Firstly, since $\mathbb{N}^2_F$'s first-order domain is ω and it provides the exact same interpretation of L^1's non-logical vocabulary as $\mathbb{N}_F$, for every L^1-formula α such that $\mathbb{N}_F \models \alpha$, $\mathbb{N}^2_F \models \alpha$. Thus, $\mathbb{N}^2_F \models A^1 \cup F$. Secondly, both the Induction Axiom and the Comprehension Schema are true in $\mathbb{N}^2_F$. This is trivial, since they were true in $\mathbb{N}^2$. The only thing that has changed by expanding this model to $\mathbb{N}^2_F$ 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, $\mathbb{N}^2_F \models A^2 \cup F$.

Now assume for *reductio* 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 set of such members. Hence, $A^2 \cup F \vdash \bot$, which is absurd, for we have shown that every $A^2 \cup F$ has a model. ■

Theorem 3.4 implies that, if infinitely many formulae are necessary for ω-inconsistency in a first-order theory, its second-order rewriting is a consistent system. 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. Hence, 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-inconsistency of a first-order theory is a product of adding (no less than) infinitely many sentences its second-order rewriting is thus consistent.

Theorem 3.3's result, on the contrary, cannot be generalized: it is not the case that any time we get ω-inconsistency by adding only a finite number of formulae to a first-order arithmetical system, the second-order rewriting of the resulting theory in inconsistent. If it were, $\mathcal{PA}^2$ would prove every arithmetically true Π^0_1-statement, which it does not.

4. Yablo's 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 to the language as a new primitive. I will stick to this second path. Our concern here is firstly with $\mathcal{YT}$'s second-order rewriting. As this system is consistent but ω-inconsistent, some results of the previous section will become handy.

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 all first-order instances of the Induction Schema. Thus, every $\mathcal{YT}$ axiom is a theorem 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 result, there are no non-intended models to guarantee its consistency, we may feel inclined to believe that this system is inconsistent *simpliciter*, like $\mathcal{PA} \cup \{\neg G\}$'s second-order rewriting.

4.2. Bad News and Good News

Unfortunately, just like the first-order calculus, the second-order calculus is not powerful enough to mirror the original informal reasoning and give 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 $\mathcal{PA} \cup \{\text{UYDP}\}$ (with all instances of the Induction Schema generated by $\mathcal{L}_{TY}$-formulae) 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 expand $\mathbb{N}$ to $\mathbb{N}_F$ in the following way:

- If $F = \varnothing$, let $Y^{\mathbb{N}_F} = T^{\mathbb{N}_F} = \varnothing$.
- If $F \neq \varnothing$, 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 $Y^{\mathbb{N}_F} = \{m\}$ and $T^{\mathbb{N}_F} = \{\text{the Gödel number of } Y(\bar{m})\}$.

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

- If $F = \varnothing$, $\mathbb{N}_F \vDash \forall x(Y(x) \leftrightarrow x \neq x) \wedge \forall x(T(x) \leftrightarrow x \neq x)$, as $Y^{\mathbb{N}_F} = T^{\mathbb{N}_F} = \varnothing$. Thus, all instances of the Induction Schema containing T or Y are true, since the ones containing $x \neq x$ instead of $Y(x)$ or $T(x)$ are true. Also, the UYDP comes out true in $\mathbb{N}_F$, just like $\forall x(T\ulcorner Y(\dot{x})\urcorner \leftrightarrow x \neq x)$, because $\mathbb{N}_F \vDash \forall x \neg T(x)$.
- If $F \neq \varnothing$, $\mathbb{N}_F \vDash \forall x(Y(x) \leftrightarrow x = \bar{m}) \wedge \forall x(T(x) \leftrightarrow x = \ulcorner Y(\bar{m})\urcorner)$, since $Y^{\mathbb{N}_F} = \{m\}$ and $T^{\mathbb{N}_F} = \{\text{the Gödel number of } Y(\bar{m})\}$. Again, all instances of the Induction Schema containing T or Y are true, for the ones containing $x = \bar{m}$ and $x = \ulcorner Y(\dot{x})\urcorner$ instead, respectively, are so.[24] $\mathbb{N}_F \vDash$ UYDP, as $\mathbb{N}_F \vDash T\ulcorner Y(\bar{m})\urcorner \wedge Y(\bar{m})$ and $\mathbb{N}_F \vDash \neg T\ulcorner Y(\bar{n})\urcorner \wedge \neg Y(\bar{n})$ for all $n \in \omega$ such that $n \neq m$. Finally, $\mathbb{N}_F \vDash F$. Let $Y(\bar{n}) \leftrightarrow \forall x(x > \bar{n} \rightarrow \neg T\ulcorner Y(\dot{x})\urcorner) \in F$. If $n = m$, both $\mathbb{N}_F \vDash Y(\bar{n})$ and $\mathbb{N}_F \vDash \forall x(x > \bar{n} \rightarrow \neg T\ulcorner Y(\dot{x})\urcorner)$, since the only member of $Y^{\mathbb{N}_F}$ is m. If, instead, $n \neq m$, then both $\mathbb{N}_F \vDash \neg Y(\bar{n})$ and $\mathbb{N}_F \vDash \neg\forall x(x > \bar{n} \rightarrow \neg T\ulcorner Y(\dot{x})\urcorner)$, for the same reasons as before.

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 the 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 Yablo's Paradox' case: any finite subset of Yablo's 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 (but reasonable) truth-theoretical principles for handling Yablo's biconditionals, even strictly second-order ones. Let A^2 be any arithmetical second-order system formulated in $\mathcal{L}^2_{TY}$, and let $\mathcal{TT}$ be any set of reasonable principles governing the truth predicate[25] 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.*

[24] 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.

[25] $\mathcal{TT}$ stands for 'Theory of Truth'.

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 inconsistency, i.e., if it is paradoxical or not in our first sense of the term—or only a finite number of Yablo's 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-theoretical principles we embrace, as long as they are reasonable, 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.*

PROOF. Straightforward, by Corollary 3.2. ■

The fact that $\mathcal{YT}^2$ lacks models implies that, assuming $\mathcal{PA}^2$ and the UYDP, there is no way of assigning truth-values consistently to Yablo's sentences —expressions of the form $Y(\bar{n})$, which state that $\forall x(x > \bar{n} \rightarrow \neg T\ulcorner Y(\dot{x})\urcorner)$. For, since $\mathcal{PA}^2 \cup \{\text{UYDP}\}$ 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 biconditional in the list and, hence, at least one of them must be false. Thus, for some such 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, Yablo's sequence is not paradoxical in either of the two senses of the term mentioned in the Introduction. Neither is it possible to derive a contradiction from reasonable arithmetical and truth-theoretical principles along with the list in the first-order calculus, nor is the formalized version of Yablo's original piece of reasoning valid in it.

The second-order calculus does not allow us to get a contradiction from reasonable arithmetical and truth-theoretical principles along with the list either. 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.[26] If we embrace the second-order notion of logical consequence we must subscribe to the idea that the second-order calculus is not powerful enough for representing Yablo's argument, and neither is the first-order calculus.

Second-order logic has an advantage over first-order logic. By regarding the intuitive reasoning as semantically valid, first, it proves $\mathcal{YT}^2$'s unsatisfiability and, thus, the impossibility of assigning truth-values consistently to Yablo's sentences. Consequently, second-order logic 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. In contrast, first-order logic cannot do any of that.

Acknowledgements. Thanks to Diego Tajer, Federico Pailos, Ignacio Ojea, Lucas Rosenblatt and Paula Teijeiro for numerous discussions and comments on earlier drafts. I would also like to thank both anonymous reviewers for helping me improve my paper. I am specially indebted to Eduardo Barrio, Roy T. Cook and Volker Halbach for useful suggestions and corrections.

References

[1] Barrio, E., Theories of truth without Standard Models and Yablo's sequences, *Studia Logica* 96:375–391, 2010.

[2] Belnap, N., and A. Gupta, *The Revision Theory of Truth*, MIT Press, Cambridge, 1993.

[3] Benacerraf, P., and C. Wright, Skolem and the Skeptic, *Proceedings of the Aristotelian Society, Supplementary Volume* 56:85–115, 1985.

[4] Bolander, T., V. F. Hendricks, and S. A. Pedersen (eds.), *Self-Reference*, CSLI Publications, Stanford, 2004.

[5] Cook, R. T., Patterns of paradox, *The Journal of Symbolic Logic* 69(1):767–774, 2004.

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

[7] Dedekind, R., 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, Oxford, 1996, pp. 787–832.

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

[26] 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.

[9] Field, H., *Saving Truth from Paradox*, Oxford University Press, New York, 2008.
[10] Forster, T., The significance of Yablo's paradox without self-reference, http://www.dpmms.cam.ac.uk/tf/, 1996.
[11] Halbach, V., *Axiomatic Theories of Truth*, Cambridge University Press, New York, 2011.
[12] Hardy, J., Is Yablo's paradox liar-like?, *Analysis* 55(3):197–198, 1995.
[13] Ketland, J., Bueno and Colyvan on Yablo's paradox, *Analysis* 64:165–172, 2004.
[14] Ketland, J., Yablo's paradox and ω-inconsistency, *Synthese* 145:295–307, 2005.
[15] Kripke, S., Outline of a theory of truth, *The Journal of Philosophy* 72:690–716, 1975.
[16] Leitgeb, H., Theories of truth which have no standard models, *Studia Logica* 68:69–87, 2001.
[17] Leitgeb, H., What is a self-referential sentence? Critical remarks on the alleged (non-)circularity of Yablos paradox, *Logique and Analyse* 177:3–14, 2002.
[18] Priest, G., The structure of the paradoxes of self-reference, *Mind* 103:25–34, 1994.
[19] Priest, G., Yablo's paradox, *Analysis* 57:236–242, 1997.
[20] Shapiro, S., *Foundations Without Foundationalism: A Case for Second-Order Logic*, Oxford University Press, New York, 1991.
[21] Sorensen, R. A., Yablo's paradox and kindred infinite liars, *Mind* 107:137–155, 1998.
[22] Tennant, N., On paradox without self-reference, *Analysis* 55:199–207, 1995.
[23] Yablo, S., Truth and reflexion, *Journal of Philosophical Logic* 14:297–349, 1985.
[24] Yablo, S., Paradox without self-reference, *Analysis* 53:251–252, 1993.
[25] Yablo, S., Circularity and paradox, in T. Bolander, V. F. Hendricks, and S. A. Pedersen (eds.), *Self-Reference*, CSLI Publications, Stanford, 2004, pp. 139–157.

Lavinia María Picollo
Department of Philosophy
University of Buenos Aires
Puan 480
Buenos Aires, Argentina
lavipicollo@gmail.com