Sometime a Paradox, Now Proof: Non-First-Order-izability of Yablo’s Paradox

SAEED SALEHI, Research Institute for Fundamental Sciences (RIFS), University of Tabriz, P.O.Box 51666-16471, Tabriz, Iran. School of Mathematics, Institute for Research in Fundamental Sciences, P.O.Box 19395–5746, Tehran, Iran. E-mail: root@SaeedSalehi.ir

Abstract
Paradoxes are interesting puzzles in philosophy and mathematics, and they could be even more fascinating when turned into proofs and theorems. For example, Liar’s paradox can be translated into a propositional tautology, and Barber’s paradox turns into a first-order tautology. Russell’s paradox, which collapsed Frege’s foundational framework, is now a classical theorem in set theory; implying that no set of all sets can exist. Paradoxes can be used in proofs of some other theorems; Liar’s paradox has been used in the classical proof of Tarski’s theorem on the undefinability of truth in sufficiently rich languages. This paradox (and also Richard’s paradox) appears implicitly in Gödel’s proof of his celebrated first incompleteness theorem. In this paper, we study Yablo’s paradox from the viewpoint of first and second order logics. We prove that a formalization of Yablo’s paradox (which is second-order in nature) is non-first-order-izable in the sense of George Boolos (1984).

2010 AMS Subject Classification: 03B05 · 03B10 · 03C07.
Keywords: Yablo’s Paradox · Non-first-orderizability.

This was sometime a paradox, but now the time gives it proof.
—William Shakespeare (Hamlet, Act 3, Scene 1).

1 Introduction
If mathematicians and philosophers have come to the conclusion that some (if not almost all) of the paradoxes cannot be (re)solved, or as Priest [11] p. 160 puts it “the programme of solving the paradoxes is doomed to failure”, they have learned to live (and also get along) with the paradoxes; of course, as long as the paradoxes do not crumble the foundations of our logical systems. Paradoxes have proved to be more than puzzles or destructive contradictions; indeed they have been used in proofs of some fundamental mathematico-logical theorems. Let us take the most well-known, and perhaps the oldest, paradox: Liar’s paradox. When translated into the language of logic, this paradox seems to claim the existence of a sentence \( \lambda \) such that \( \lambda \leftrightarrow \neg \lambda \) holds. Now, Liar’s paradox can turn into a propositional tautology: \( \neg (\rho \leftrightarrow \neg \rho) \). In fact, when trying to convince oneself about the truth of \( \neg (\rho \leftrightarrow \neg \rho) \), one can see that the supposed argument is not that much different from the argument of Liar’s paradox. One can clearly see that the paradox becomes a (semantic) proof for that tautology; hence the title of the article (which uses the above epigram of Shakespeare).
Let us take a second example; Russell’s paradox. If there existed a set \( r \) such that \( \forall x (x \in r \iff x \notin x) \), then we would have a contradiction (the sentence \( r \subseteq r \iff r \notin r \) which results from substituting \( x \) with \( r \)). So, the sentence \( \neg \exists y \forall x (x \in y \iff x \notin x) \) is a theorem in the theory of sets, whose proof is nothing more than the argument of Russell’s paradox. Going deeper into the proof (or the paradox), one can see that no real set-theoretic properties of the membership relation (\( \in \)) is used. That is to say that for an arbitrary binary relation \( s \), the sentence \( \neg \exists y \forall x [s(y, x) \iff \neg s(x, x)] \) is a first-order logical tautology (see [11, Exercise 12, p. 76]). Now, if we interpret the predicate \( s(y, x) \) as “\( y \) shaves \( x \)”, then we get Barber’s paradox (due to Russell again). More generally, for any formula \( \varphi(x, y) \) with the only free variables \( x \) and \( y \), the sentence \( \neg \exists y \forall x [\varphi(x, y) \iff \neg \varphi(x, x)] \) is a first-order logical tautology, whose semantic proof is very similar to the argument of Russell’s or Barber’s paradox. In a similar way, \( \neg \exists X^{(2)} \exists y \forall x [X^{(2)}(x, y) \iff \neg X^{(2)}(x, x)] \) is a second-order tautology.

In this paper, we are mainly interested in Yablo’s paradox [13, 14]; several papers (that we do not cite all of them here) and one book [4] have been written on different aspects of this paradox. Yablo’s paradox says that if there existed a sequence of sentences \( \{Y_n\}_{n \in \mathbb{N}} \) with the property that for all \( n \in \mathbb{N} \), the sentence \( Y_n \) is true if and only if \( Y_k \) is untrue for every \( k > n \), then we would have a contradiction, since none of those sentences can have a truth value (the sentences \( Y_n \) would be neither true nor false). This paradox is humbly called by Yablo himself, the \( \omega \)-Liar paradox. The paradoxicality of the sequence \( \{Y_n\}_{n \in \mathbb{N}} \) of sentences with the above property follows form the observation that if \( Y_m \) is true, for some \( m \), then \( Y_{m+1} \), and also all \( Y_k \)’s, for \( k > m+1 \), should be untrue. So, by the falsity of \( Y_{m+1} \), there should exist some \( j > m+1 \) such that \( Y_j \) is true; a contradiction. Whence, all \( Y_m \)’s should be untrue, and so \( Y_0 \) must be true; another contradiction!

## 2 Yablo’s Paradox in Second-Order Logic

For formalizing Yablo’s paradox in a (first-order or second-order) language, we first abstract away even the order relation, that appears in the paradox, and replace it with an arbitrary binary relation symbol \( R \); see [6] for a non-arithmetical formulation of Yablo’s paradox. Let us take \( Y_1 \) to be the first-order scheme

\[
\neg \forall x (\varphi(x) \iff \forall y [x R y \rightarrow \neg \varphi(y)]),
\]

where \( \varphi(x) \) is an arbitrary first-order formula with the only free variable \( x \). Here, the sentences \( Y_n \) are represented by \( \varphi(n) \), and the quantifiers of the form \( \forall k > n \cdot \cdot \cdot \) are represented by \( \forall k (k R n \rightarrow \cdot \cdot \cdot) \).

**Definition 2.1** (\( Y \): Yablo’s Paradox in Second Order Logic)

Let \( Y \) be the following second-order sentence:

\[
\neg \exists X^{(1)} \forall x [X^{(1)}(x) \iff \forall y [x R y \rightarrow \neg X^{(1)}(y)]],
\]

where \( R \) is a fixed binary relation symbol.\( \diamond \)

Some sufficient conditions for proving (\( Y_1 \) and) \( Y \) are

* A closely related paradox is Visser’s [12] which we do not study here.
(A1): $\forall x \exists y (xRy)$, and (A2): $\forall x, y, z (xRy \land yRz \rightarrow xRz)$.

That is to say that (A1 $\land$ A2) $\rightarrow \mathcal{Y}_1$ is a first-order tautology, and) A1 $\land$ A2 $\rightarrow \mathcal{Y}$ is a second-order tautology; see [9, 10]. None of these conditions are necessary for $\mathcal{Y}$; for example in the directed graph $\langle D; R \rangle$ with $D = \{a, b, c\}$ and $R = \{(a, b), (a, c), (c, c)\}$, we have $\mathcal{Y}$ and $A_2$ but not $A_1$. Also, in the directed graph $\langle D; R \rangle$ with $D = \{a, b, c\}$ and $R = \{(a, b), (b, c), (c, c)\}$, we have $\mathcal{Y}$ and $A_1$ but not $A_2$.

As a matter of fact, some weaker conditions than A1 $\land$ A2 can also prove ($\mathcal{Y}_1$ and) $\mathcal{Y}$. For example, the sentence

$$ (A): \forall x \exists y (xRy \land \forall z [yRz \rightarrow xRz]), $$

suffices (see Theorem 2.3 below). To see that A is really weaker than A1 $\land$ A2, consider $\langle D; R \rangle$ with $D = \{a, b, c, d\}$ and $R = \{(a, b), (b, c), (a, d), (b, d), (c, d), (d, d)\}$. Then $\langle D; R \rangle$ does not satisfy A2, since it is not transitive (we have $aRbRc$ but $\neg aRc$), while it satisfies A, since for any $x \in D$ we have $xRd \land \forall z [dRz \rightarrow xRz]$. Even some weaker conditions than A can prove ($\mathcal{Y}_1$ and) $\mathcal{Y}$.

**Definition 2.2 (Some Sufficient Conditions for Proving $\mathcal{Y}$)**

Let $\theta_0(x)$ be the formula $\exists y (xRy \land \forall z [yRz \rightarrow xRz])$.

For any $n \in \mathbb{N}$, if $\theta_n(x)$ is defined, then let $\theta_{n+1}(x) = \exists y (xRy \land \forall z [yRz \rightarrow \theta_n(z)])$.

We now show that $\{\forall x \theta_n(x)\}_{n \in \mathbb{N}}$ is a decreasing sequence of sentences (i.e., every sentence is stronger than its successor, in the sense that the sentence implies its successor but not vice versa) that all imply ($\mathcal{Y}_1$ and) $\mathcal{Y}$. Note that $A = \forall x \theta_0(x)$.

**Theorem 2.3** ($\forall x \theta_0(x) \not\vDash \cdots \not\vDash \forall x \theta_n(x) \not\vDash \forall x \theta_{n+1}(x) \not\vDash \cdots \not\vDash \mathcal{Y}$)

For every $n \in \mathbb{N}$, we have

1. $\forall x \theta_n(x) \not\vDash \mathcal{Y}$;
2. $\forall x \theta_n(x) \vDash \forall x \theta_{n+1}(x)$;
3. $\forall x \theta_{n+1}(x) \not\vDash \forall x \theta_n(x)$.

**Proof:**

(1): By induction on $n$. For $n = 0$, take a directed graph $\langle D; R \rangle$ and assume that $\forall x \theta_0(x)$ holds in it. If a subset $X \subseteq D$ satisfies $\forall x (x \in X \leftrightarrow \forall y [xRy \land \neg y \in X])$, then for any $a \in D$, the assumption $a \in X$ implies that there exists some $b \in D$ such that $aRb$ and $\forall z [bRz \rightarrow aRz]$. Now, by $b \not\in X$, there should exist some $c \in D$ such that $bRc$ and $c \not\in X$. Also, $aRc$ should hold, which is a contradiction with $a \in X$. Thus $X = \emptyset$. But then for any $a \in D$ there should exist some $b \in D$ with $aRb$ and $b \in X$, and so $X \neq \emptyset$; another contradiction. Thus, there exists no such $X \subseteq D$; whence, $\forall x \theta_0(x) \not\vDash \mathcal{Y}$.

Now, suppose that $\forall x \theta_n(x) \not\vDash \mathcal{Y}$ holds. Take a directed graph $\langle D; R \rangle$ and assume that $\forall x \theta_{n+1}(x)$ holds in it. If for some $X \subseteq D$ we have $\forall x (x \in X \leftrightarrow \forall y [xRy \land \neg y \in X])$, then for any $a \in D$, there exists some $b \in D$ such that $aRb$, and we have $aRx$ for any $x$ in the set $D_b = \{z \in D \mid bRz\}$. Now, if $D_b \neq \emptyset$, then the directed graph $\langle D_b, R \cap D_b^2 \rangle$ satisfies $\forall x \theta_n(x)$, and so, by the induction hypothesis, the set $X \cap D_b$ cannot exist. So, we necessarily have $D_b = \emptyset$. Now, if $a \not\in X$ holds, then we should have that $b \not\in X$ and so there exists some $c \in D_b$ with $c \not\in X$; a contradiction. Thus, $X = \emptyset$. Then, for any $a \in D$, since $a \not\in X$, there should exist some $b$ with $aRb$ and $b \in X$; another contradiction. This shows that $\forall x \theta_{n+1}(x) \not\vDash \mathcal{Y}$.

(2): Suppose that $\forall x \theta_n(x)$ holds, and fix an $x$. There is some $y$ such that $xRy$; and for any $y$ with $yRz$ we have $\theta_n(z)$ by the assumption $\forall x \theta_n(x)$. So, $\forall x \theta_{n+1}(x)$ holds.

(3): Consider $\langle D; R \rangle$, with $D = \{a_0, a_1, \ldots, a_{2n}\}$ and $R = \{(a_i, a_{i+1}) \mid 0 \leq i < 2n\} \cup \{(a_{2n}, a_{2n})\}$. In the directed graph $\langle D; R \rangle$, obviously, $\forall x \theta_{n+1}(x)$ holds, but $\forall x \theta_n(x)$ does not hold, since we have $a_{2n-2}Ra_{2n-1}Ra_{2n}$ but $\neg (a_{2n-2}Ra_{2n})$.

\[\square\]
4 Non-First-Orderizability of Yablo’s Paradox / Saeed Salehi

As a result, \( \mathcal{Y} \) does not imply the sentence \( \forall x \theta_n(x) \), for any \( n \in \mathbb{N} \). In the next section, we show that no first-order sentence in the language of \( \langle R \rangle \) is equivalent with the second-order sentence \( \mathcal{Y} \). So, neither the sentence \( \mathcal{Y} \) nor its negation \( \neg \mathcal{Y} \) is first-order-izable (see [2; 3]). Not only the second-order sentence \( \mathcal{Y} \) is non-equivalent with any first-order sentence, but also it is non-equivalent with any first-order theory (which could contain infinitely many sentences). Actually, \( \neg \mathcal{Y} \) is equivalent with the existence of a kernel in a directed graph \( \langle D; R \rangle \); see e.g. [4]. So, our result shows that the existence or non-existence of a kernel in a directed graph is not equivalent to any first-order sentence (in the language of directed graphs). Whence, Yablo’s paradox, formalized as \( \langle Y_1 \text{ or as} \rangle \mathcal{Y} \) in Definition 2.1 does not turn by itself into (a first-order or) a second-order tautology, and some conditions should be put on \( R \) to make it a theorem. This paradox can be nicely translated into some theorems in Linear Temporal Logic (see [8; 7]) or in Modal Logic (see [5]).

3 Non-first-orderizability of Yablo’s Paradox

Consider the language \( \langle \mathcal{s} \rangle \), where \( \mathcal{s} \) is a unary function symbol. A standard structure on this language is \( \langle \mathbb{N}; \mathcal{s} \rangle \), where \( \mathcal{s} \) is interpreted as the successor function: \( \mathcal{s}(n) = n + 1 \) for all natural numbers \( n \in \mathbb{N} \).

**Definition 3.1 (Theory of Successor, and Kernel of a Directed Graph)**

Let the theory \( \mathcal{S} \) on the language \( \langle \mathcal{s} \rangle \) consist of the following axiom:

\[
\forall x, y \left[ \mathcal{s}(x) = \mathcal{s}(y) \rightarrow x = y \right].
\]

With any structure \( \langle M; \mathcal{s} \rangle \), the directed graph \( \langle M; R \rangle \) is associated, where \( R \) is defined by \( xRy \iff y = \mathcal{s}(x) \), for all \( x, y \in M \).

For a directed graph \( \langle D; R \rangle \), a subset \( K \subseteq D \) is called a Kernel, when it has the following property: \( \forall x \left( x \in K \iff \forall y \left[ xRy \rightarrow y \notin K \right] \right) \).

For a formula \( \varphi \) over the language \( \langle \mathcal{R} \rangle \), let \( \varphi^\mathcal{s} \) result from \( \varphi \) by replacing each \( uRv \) with \( \mathcal{s}(u) = v \) for variables \( u, v \); so, \( \varphi^\mathcal{s} \) is a formula over the language \( \langle \mathcal{s} \rangle \).

So, \( \neg \mathcal{Y} \) states the existence of a Kernel in a directed graph with relation \( R \), and the \( \mathcal{s} \)-translation of Yablo’s paradox \( \mathcal{Y}^\mathcal{s} \) is equivalent with the second-order sentence \( \neg \exists X(\forall x \left[ X(x) \iff \neg X(\mathcal{s}(x)) \right]) \). Any structure \( \langle M; \mathcal{s} \rangle \) which satisfies \( \mathcal{S} \) may contain some copies of

\[
\mathbb{N} \approx \{ a_0, a_1, a_2, \ldots \}
\]

with \( a_{n+1} = \mathcal{s}(a_n) \) for all \( n \in \mathbb{N} \), such that there is no \( a \in M \) with \( \mathcal{s}(a) = a_0 \). It may also have some copies of

\[
\mathbb{Z} \approx \{ \ldots, a_{-2}, a_{-1}, a_0, a_1, a_2, \ldots \},
\]

in which \( a_{m+1} = \mathcal{s}(a_m) \) for all \( m \in \mathbb{Z} \). There could be also some finite cycles

\[
\mathbb{Z}_m \approx \{ a, \mathcal{s}(a), \mathcal{s}^2(a), \ldots, \mathcal{s}^{m-1}(a) \} \quad \text{with} \quad \mathcal{s}^m(a) = a,
\]

for some \( m > 0 \). Let us note that, by the axiom \( \mathcal{S} \), no two copies of \( \mathbb{N} \) or \( \mathbb{Z} \) or a finite cycle can intersect one another. Indeed, these are all a model \( \langle M; \mathcal{s} \rangle \) of \( \mathcal{S} \) can contain.
Lemma 3.2 (Axiomatizability of $\neg \psi + S$)
The associated directed graph $(D; R)$ of a model $(M; s)$ of $S$ satisfies $\neg \psi$, if and only if $(M; s)$ has no odd cycles, if and only if the structure $(M; s)$ also satisfies the axioms $\neg \exists x(\psi^{2n+1}(x) = x)$ for $n \in \mathbb{N}$.

Proof:
The second equivalence is straightforward; so, we prove the first equivalence only.
First, suppose that $(M; s)$ has no odd cycles. Then let $K \subseteq M$ consist of the even natural and integer numbers (as the copies of $\mathbb{N}$ and $\mathbb{Z}$) of $M$ (if any), and the elements with even indices in the finite cycles of $M$; i.e., for a finite even cycle $\{a, s(a), s^2(a), \cdots, s^{2m+2}(a) = a\}$ take $\{a, s(a), s^4(a), \cdots, s^{2m}(a)\}$. Then the set $K$ is a kernel of $(M; R)$, since an element of $M$ is in $K$, if and only if it is even indexed, and only if its successor is odd indexed, if and only if its successor is not in $K$.
Thus, $(M; R)$ satisfies $\neg \psi$. This would have not been possible if there were an odd cycle; i.e., an element $\alpha$ such that $s^{2m+1}(\alpha) = \alpha$ for some $m > 0$, since $\alpha$ would have been odd and even indexed at the same time.
Second, suppose that the directed graph $(M; R)$ associated to $(M; s)$ has a kernel $K$, and also (for the sake of a contradiction) that $(M; s)$ has an odd cycle such as $\{a, s(a), s^2(a), \cdots, s^{2m+1}(a) = a\}$, for some $m > 0$. Then, if $a \in K$, then $s(a) \notin K$, then $\cdots$ $s^{2m}(a) \in K$, and so $a = s^{2m+1}(a) \notin K$, a contradiction. Also, if $a \notin K$, then $s(a) \in K$, then $\cdots$ $s^{2m}(a) \notin K$, and so $a = s^{2m+1}(a) \in K$, a contradiction again. Therefore, if $(M; R)$ has a kernel, then $(M; s)$ can have no odd cycle. 

Theorem 3.3 (Non-First-Order-izability of $\psi$ and $\neg \psi$)
The second-order sentence $\psi$ is not equivalent with any first-order sentence.

Proof:
If there were a first-order sentence in the language $(R)$ equivalent to $\psi$, then by Lemma 3.2, $S' = S \cup \neg \exists x(\psi^{2n+1}(x) = x) | n \in \mathbb{N}$ would be finitely axiomatizable (see [11, Lemma 4.2.9]). But this is not true, since for any finite subset of this theory, there exists a structure that satisfies that finite sub-theory but is not a model of the whole theory $S'$; it suffices to take a sufficiently large odd cycle.

Thus, there can exist no first-order sentence $\eta$ such that the second-order sentence $\eta \leftrightarrow \psi$ is a logical tautology. As a result, the proposed formalization $\psi$ of Yablo's paradox in Definition 2.1, being second-order in nature, is not (equivalent with any) first-order (sentence). We end the paper with a stronger result: there cannot exist any first-order theory that is equivalent with $\psi$. So, Yablo's paradox is not even infinitely first-order (i.e., it is non-equivalent even with any theory that consists of an infinite set of first-order sentences).

Theorem 3.4 (Non-Equivalence of $\psi$ With First-Order Theories)
The second-order sentence $\psi$ is not equivalent with any first-order theory.

Proof:
By Lemma 3.2, the theory $\neg \psi + S$ is axiomatizable over $(s)$; if $\psi$ were axiomatizable, then by [11, Lemma 4.2.10] the theory $\neg \psi + S$ would be finitely axiomatizable. But this contradicts Theorem 3.4. So, $\psi$ is not axiomatizable, hence $\psi$ is not equivalent with any first-order theory.
We conjecture that the second-order sentence \(\neg Y\), too, is non-equivalent with all the first-order theories. This does not concern the main topic of this article, since the sentence \(\neg Y\) does not express Yablo’s paradox, while \(Y\) does that.

Acknowledgements
This research was partially supported by a grant from IPM (No 95030033). I warmly thank Kaave Lajevardi for drawing my attention to the Shakespearean epigram, that is quoted in this paper.

References
[1] Berge, Claude; Graphs, North Holland (3rd ed. 1991). isbn: 9780444876034
[2] Boolos, George; To Be is To Be a Value of a Variable (or to be some values of some variables), The Journal of Philosophy 81:8 (1984) 430–449. doi: 10.2307/2026308 Reprinted in: Boolos, G.; Logic, Logic and Logic (isbn: 9780674537668) Harvard University Press (1998) pp. 54–72.
[3] Boolos, George; Nonfirstorderizability Again, Linguistic Inquiry 15:2 (1984) p. 343. https://www.jstor.org/stable/4178386
[4] Cook, Roy; The Yablo Paradox: An Essay on Circularity, Oxford University Press (2014). isbn: 9780199669608
[5] Forster, Thomas & Gor é, Rajeev; Yablo’s Paradox as a Theorem of Modal Logic, Logique et Analyse 59:235 (2016) 283–300. doi: 10.2143/LEA.235.0.3170108
[6] Halbach, Volker & Zhang, Shuoying; Yablo Without Gödel, Analysis 77:15 (2017) 53–59. doi: 10.1093/analys/anw062
[7] Karimi, Ahmad & Salehi, Saeed; Diagonal Arguments and Fixed Points, Bulletin of the Iranian Mathematical Society 43:5 (2017) 1073–1088. https://bit.ly/31pQVYF
[8] Karimi, Ahmad & Salehi, Saeed; Theoremizing Yablo’s Paradox, arXiv:1406.0124 (2014) 7 pp.
[9] Ketland, Jeffrey; Yablo’s Paradox and \(\omega\)-Inconsistency, Synthese 145:3 (2005) 295–302. doi: 10.1007/s11229-005-6201-6
[10] Priest, Graham; Unstable Solutions to the Liar Paradox, in: S. J. Bartlett & P. Suber (eds.), Self-Reference: Reflections on Reflexivity (isbn: 9789024734740), Martinus Nijhoff Publishers (1987) pp. 145–175. doi: 10.1007/978-94-009-3551-8_9
[11] van Dalen, Dirk; Logic and Structure, Springer (5th ed. 2013). isbn: 9781447145578
[12] Visser, Albert; Semantics and the Liar Paradox, in: D. Gabbay & F. Günthner (eds.), Handbook of Philosophical Logic, Volume IV: Topics in the Philosophy of Language (isbn: 9789401070218), Reidel (1989) pp. 617–706. doi: 10.1007/978-94-009-1171-0_10
[13] Yablo, Stephen; Truth and Reflection, Journal of Philosophical Logic 14:3 (1985) 297–349. doi: 10.1007/BF00249368
[14] Yablo, Stephen; Paradox Without Self-Reference, Analysis 53:4 (1993) 251–252. doi: 10.1093/analys/53.4.251