Yablo's paradox and ω-inconsistency

Shunsuke Yatabe
Graduate School of Letters, Kyoto University

We focus on the *co-inductive* character of Yablo's paradox, analyzing it by comparison in truth theories and in **ZFA**. We show that the ω-inconsistency of truth theories is because, while they allow mixtures of induction and co-induction, such mixtures are impossible in an ω-consistent **ZFA**.

Introduction

Let us assume there exist infinitely many propositions $\langle S_0, S_1, S_2, \cdots \rangle$ such that S_n insists that S_i is false for any $i > n$. Then these propositions imply a contradiction. First let us assume S_0 is false. Then there must be $j > 0$ such that S_j is true. This means that all $S_{j+1}, S_{j+2}, S_{j+3}, \cdots, S_k, \cdots$ must be false. However, if S_{j+1} is false, then there exists $k > j+1$ such that S_k is true, a contradiction. Next assume S_0 is true. Then $S_1, S_2, \cdots$ are false, identical to the previous case. This is the well-known Yablo's paradox [Yab93].

There has been previous discussion as to whether Yablo's paradox is self-referential, but this paper does not address that topic. Instead, we focus on how Yablo propositions are constructed. The answer to whether Yablo's paradox is self-referential seems to depend on how these propositions are constructed, and the essence of their construction. A source of trouble is that only one method of constructing $\langle S_n : n \in \omega \rangle$ is known, using diagonalization in a truth theory in classical logic [P97]. Such deficiency of comparison examples could lead to a mistake that we regard properties that contingently hold in the truth theory (and do not hold in other theories) as essential properties of Yablo's paradox. For example, consistent truth theories with sufficient expressive power to define Yablo propositions should be ω-inconsistent [L01], but we do not know whether ω-inconsistency is essential in Yablo's paradox.

Yablo propositions satisfy a characteristic property in that the intuitive meaning of S_i is $\bigwedge_{j>i} \neg\mathbf{Tr}(\ulcorner S_j \urcorner)$ (if the language has an infinite conjunction). Therefore,

$$
\begin{aligned}
S_i &\equiv \neg\mathbf{Tr}(\ulcorner S_{i+1} \urcorner) \wedge S_{i+2} \\
\neg S_i &\equiv \mathbf{Tr}(\ulcorner S_{i+1} \urcorner) \vee \neg S_{i+2}.
\end{aligned}
$$

This means each S_i is constructed by directly using S_{i+1} and S_{i+2}. However, to construct S_{i+2}, we need S_{i+3} and S_{i+4}, etc. In this way, there is an infinite regress; we need infinitely many $\langle S_{i+1}, S_{i+2}, S_{i+3}, \cdots \rangle$ to construct S_i in the end. The characteristic points of this construction are that (1) we only directly use finitely many already-constructed objects to construct a new object, and (2) we need infinitely many steps to reach the *initial* construction case (meaning this is not inductive construction).

Such constructions are called *co-inductive*, and are widely used in computer science to represent behaviors of non-terminate automatons [C93] because they allow construction of *potentially* infinite objects in a finite way. Yablo's paradox seems to be evidence that co-induction is naturally used in natural language. We focus on the *co-inductive* character of Yablo's paradox, and analyze it by comparing the paradox in truth theories to that in **ZFA**.

Preliminaries on ZFA

One of the most famous ways to define a co-inductive language, a language with co-inductively defined formulae, is to use **ZFA** [BE87] [BM96]. This is done by coding co-inductively defined propositions by hypersets. As for Yablo's paradox, Yablo suggested fixing **ZFA** as an analysis framework [Yab06], but abandoned this approach without serious consideration. **ZFA** is an axiomatic set theory, **ZF** minus the axiom of foundation plus the anti-foundation axiom (**AFA**), which allows definition of *hypersets*, which need not be well founded in classical logic. Due to space limitations, we present the so-called *flat system lemma* with only a brief review.

Definition 1. A flat system of equations $\langle X, A, e\rangle$ has the following characteristics:

- $X \subseteq U$ (urelements, interpreted as variables),
- A is an arbitrary set, and
- $e : X \to \mathcal{P}(X \cup A)$.

An example of a flat system is $\langle\{a\}, \emptyset, \{\langle a, \{a\}\rangle\}\rangle$ for some urelement a; since $e(a) = \{a\}$, this system represents an equation $x = \{x\}$, where x is a free variable.

Theorem 3. **AFA** guarantees that *any flat system of equations defines hypersets uniquely.*

As a sort of co-inductive definition[10], consider the flat system

$$\langle\{a_n : n \in \omega\}, \emptyset, \{\langle a_n, \{a_{n+1}, a_{n+2}\}\rangle : n \in \omega\}\rangle,$$

which represents equations $x_n = \{x_{n+1}, x_{n+2}\}$ for any n (the construction is finite in any successor step but we cannot achieve this in the initial case).

We fix **ZFA** as the framework of this paper because, thanks to [BE87], it is one of the most famous truth theory frameworks that enables purely co-inductive construction of formulae[11]. The framework of [BE87] seems to be *overkill* for semantic paradoxes. The liar proposition can be represented even as arithmetic, but **ZFA** produces hypersets, as many as ordinal well-founded sets, to represent such paradoxical propositions. The real value of this framework is that it allows many kinds of co-inductive construction[12].

Coding Yablo propositions by hypersets

Let us introduce the construction of *Russellian propositions* or *Austinian types*[13]. Define their *co-inductive* coding method by hypersets as follows:

---

[10] Actually **ZFA** is a set theory whose sets are constructed by co-induction in some transfinite induction step. The universe of **ZFA** is constructed by $\mathbf{V}_0 = \emptyset$, $\mathbf{V}_{\alpha+1} = \mathbf{V}_\alpha \bigcup \mathcal{P}^*(\mathbf{V}_\alpha)$ and $\mathbf{V}_\gamma = \bigcup_{\delta<\gamma} \mathbf{V}_\delta$ for any γ limit, where $\mathcal{P}^*(A) = \{x :\in |_{\mathrm{TC}(x)}$ is bisimilar to R for some $R \subseteq \mathbf{TC}(A)^2\}$[V04].

[11] Many theories allow co-inductive object definitions. For example, an intuitionistic theory has been extended to allow such definitions (we do not have to worry about overly rich ontologies in such theories) [C93], and naive set theories in non-classical logics have strong co-inductive characters [Yat12a]. However, **ZFA** is the most well known among them.

[12] As Yablo pointed out in [Yab06], there is a counterintuitive problem that any propositions S_i, S_j of Yablo's paradox are mutually identical. If we fix an Austinian-like approach, all propositions are pairwise distinct (*situations* are taken into consideration). We omit the details here due to space limitations.

[13] Roughly speaking, an Austinian proposition is a pair of a *situation* and an Austinian type: different definitions of situations give different definitions of propositions.

Definition 2 (Russellian propositions or Austinian types). Formulae are coinductively coded in **ZFA** as follows:

- $\ulcorner A \wedge B \urcorner = \{\{\mathbf{c}, \ulcorner A \urcorner\}, \{\mathbf{c}, \ulcorner B \urcorner\}\}$ and $\ulcorner \wedge_{i \in I} A_i \urcorner = \{\{\mathbf{c}, \ulcorner A_i \urcorner\} : i \in I\}$,
- $\ulcorner A \vee B \urcorner = \{\{\mathbf{d}, \ulcorner A \urcorner\}, \{\mathbf{d}, \ulcorner B \urcorner\}\}$ and $\ulcorner \vee_{i \in I} A_i \urcorner = \{\{\mathbf{d}, \ulcorner A_i \urcorner\} : i \in I\}$,
- $\ulcorner \neg A \urcorner = \{\mathbf{n}, \ulcorner A \urcorner\}$,
- $\ulcorner \mathbf{Tr}(A) \urcorner = \{\mathbf{t}, \ulcorner A \urcorner\}$

for some fixed set $\mathbf{c}, \mathbf{d}, \mathbf{n}, \mathbf{t}$ which are not equal to any natural numbers.

Note that this coding does not have an initial case, but is sufficient to code the liar propositions or Yablo propositions. For example, the liar proposition λ is coded by a Russellian proposition $\ulcorner \lambda \urcorner$ satisfying $x = \{0, \{*, x\}\}$.

Next let us define Yablo propositions[14].

Definition 3 (Yablo propositions). Yablo (Russellian) propositions $\{S_n : n \in \omega\}$ are coded by the following equation: let $\langle \{x_n, p_n : n \in \omega\}, \{\mathbf{c}, \mathbf{n}, \mathbf{t}\}, e \rangle$ be an infinite flat system such that, for any $n \in \omega$,

$$\begin{aligned} e(x_n) &= \{p_k : k > n\} \\ e(p_n) &= \{\mathbf{c}, q_n\} \\ e(q_n) &= \{\mathbf{n}, r_n\} \\ e(r_n) &= \{\mathbf{t}, x_k\} \end{aligned}$$

Then $S_0, S_1, \cdots$ are solutions of $x_0, x_1, \cdots$.

Theorem 4. Yablo (Russellian) propositions $\langle S_n : n \in \omega \rangle$ exists in **ZFA**[15].

The proof is a simple application of theorem 3. We note that, as we pointed out, $S_i = S_j$ holds for any i, j since there is a *bisimulation* among all $\in$-graphs $\langle S_n : n \in \omega \rangle$ by this coding[16][17]. However,

[14] We do not need the truth predicate to construct Yablo propositions in this framework. $\langle Y_n : n \in \omega \rangle$ are defined by $Y_0, Y_1, \cdots$ are solutions of $x_0, x_1, \cdots$ and $\neg Y_0, \neg Y_1, \cdots$ are solutions of $y_0, y_1, \cdots$ appearing in

$$\begin{aligned} e(p_n) &= \{\mathbf{c}, y_n\} \\ e(q_n) &= \{\mathbf{c}, x_n\} \\ e(x_n) &= \{p_k : k > n\} \\ e(y_n) &= \{q_k : k > n\} \end{aligned}$$

The intuitive meaning of Y_n is $\wedge_{n<i} \neg Y_i$, and this is equivalent to $\neg Y_{i+1} \wedge Y_{i+2}$. Recall that the liar paradox is not unique but an instance of a self-referential paradox; a Russell paradox is another. In this sense, Yablo's paradox is just an instance of a *co-inductive* paradox.

[15] Yablo's paradox implies a contradiction when applying Russellian semantics. If we apply Austinian-like semantics, all Yablo propositions are simply false (and thus do not imply a contradiction)

[16] Let us consider the meaning of this. In the paradox, first we take S_0 and assume it is true (or false). However, even though we first assume S_i is true (false), the *behavior* of the paradox, the derivation of the inconsistency, is an identical form. If we formalize the paradox using game semantics, the player who gives a counterexample has a very simple winning strategy regardless of the opponent's choice, S_0 or S_i. In this sense, S_0 and S_i are identical. Of course, the difference in the starting point can be distinguished if we consider the hidden parameter, *situations*: we can distinguish S_0 and S_i in Austinian-like Semantics.

[17] Note that the mutual equality of Yablo propositions collapses Yablo's paradox to a simple liar-like self-referential paradox. Actually, since $S_0 = S_i = S$, the paradox, $S_0 \to \neg S_i \wedge S_i$ and $\neg S_0 \to S_i \wedge \neg S_i$, are just equal to $S \to \neg S$ and $\neg S \to S$.

this is just a technical problem: just adding indexes makes them mutually different hypersets [Yat12b] (but we omit the detail because they are essentially the same). We also note that any S_n forms an infinite-branching tree of infinite height[18].

A comparison to truth theories: a source of ω-consistency

As discussed above, well-known consistent theories with sufficient expressive power, like Γ [Mc85] and $\mathbf{CT}_\omega$ [HH05], are ω-inconsistent. Yablo propositions $\langle \bar{S}_x : x \in \omega \rangle$ are constructed by the fixed point lemma in such theories as follows:

$$\bar{S}_x \quad \equiv \quad (\forall z)[z > x \to \neg \mathbf{Sat}(\ulcorner S_x \urcorner, z)],$$

where $\mathbf{Sat}(\ulcorner \varphi(x) \urcorner, z) \equiv \mathbf{Tr}(\ulcorner \varphi(z) \urcorner)$. Roughly speaking, the intuitive meaning of $\bar{S}_x$ is

$$\bar{S}_x \equiv \underbrace{\cdots \wedge \neg \mathbf{Sat}(\ulcorner \bar{S}_x \urcorner, x+2) \wedge \neg \mathbf{Sat}(\ulcorner \bar{S}_x \urcorner, x+1)}_{\infty \text{ many}}$$

The main difference between this construction and that of **ZFA** is whether the construction has an *initial* case or not. In the **ZFA** case, the construction does not have an initial case. Truth theory constructions do have an initial case S_x, however, and any S_y is constructed from S_x as $\mathbf{Sat}(\ulcorner S_x \urcorner, y)$ for any $y > x$. Thanks to the truth predicate, the fixed point lemma enables an infinite operation over formulae (S_x itself is a limit of infinite operation $\wedge_{y>x} \neg \mathbf{Sat}(\ulcorner S_x \urcorner, y)$). The construction of S_x is not by pure co-induction, but by *a mixture of induction and co-induction*, that is, a co-inductive construction with the initial case.

This mixture plays a key role in the proof of ω-inconsistency in truth theories. For example, in Γ [Mc85], ω-inconsistency is proved by the following formula γ:

$$\begin{array}{rcl} \gamma & \equiv & \neg \forall x \mathbf{Tr}(f(x, \ulcorner \gamma \urcorner)) \\ f(n, \ulcorner \varphi \urcorner) & = & \ulcorner \underbrace{\mathbf{Tr}(\ulcorner \cdots \mathbf{Tr}}_{n \text{ times}}(\ulcorner \varphi \urcorner) \cdots) \urcorner \end{array}$$

Roughly speaking, γ is defined by *a mixture of induction and co-induction* in the sense that the intuitive meaning of γ is $\gamma \equiv \neg \underbrace{\mathbf{Tr}(\ulcorner \cdots (\ulcorner \mathbf{Tr}}_{\infty \text{ many}}(\ulcorner \gamma \urcorner) \urcorner) \cdots) \urcorner)$.

Summing up, **ZFA** is proof-theoretically strong, so **ZFA** can distinguish the well-founded (WF) and non-WF parts of the universe. The set of natural numbers ω, which is a member of the WF part, is constructed by induction only, and co-inductive objects are in another partition, that is, the non-WF part. Therefore co-inductive construction does not give any effect to ω. In truth theories, the model domain only consists of natural numbers, which are constructed inductively. Co-inductive construction, which is possible by the fixed point lemma and the truth predicate, is not possible without induction, and their mixture seems to involve the existence of non-standard natural numbers.

Bibliography

[BE87] Jon Barwise. John Etchemendy. The Liar: An Essay in Truth and Circularity. Oxford University Press (1987)

[18] Each tree S_n is self-similar, i.e., for any branch t of S_n, there is a sub-tree $T \subseteq S_n|_t$ such that there is an isomorphism $\pi_j : T \to S_j$ for some $j > n$. Actually, the tree and isomorphisms form a completely iterative algebra [Mo08].

[BM96] Jon Barwise, Lawrence Moss. Vicious Circles. CSLI publications (1996)

[C93] Thierry Coquand. Infinite Objects in Type Theory. TYPES (1993) 62-78.

[HH05] Volker Halbach and Leon Horsten. The deflationist's axioms for truth. in J.C. Beal and B. Armour-Garb (eds.) Deflationism and Paradox. Oxford Universoty Press.

[L01] Hannes Leitgeb. Theories of truth which have no standard models. Studia Logica, 68 (2001) 69-87.

[Mc85] Vann McGee. How truthlike can a predicate be? A negative result. Journal of Philosophical Logic, 17 (1985): 399-410.

[Mo08] Lawrence S. Moss. Coalgebra and Circularity. Preprint (2008)

[P97] Graham Priest. Yablo's paradox. Analysis 57 (1997) pp.236-242.

[V04] Matteo Viale. The cumulative hierarchy and the constructible universe of ZFA. Mathematical Logic Quarterly, 2004, volume 50, no. 1, pp.99-103.

[Yab93] Stephen Yablo. Paradox Without Self-Reference. Analysis 53 (1993) 2251-52.

[Yab06] Stephen Yablo. Circularity and Paradox. Self-reference. in Bolander, Hendricks, Pedersen (eds.), Self-reference, Stanford: CSLI Publications.

[Yat12a] Shunsuke Yatabe. A constructive naive set theory and infinity. Preprint.

[Yat12b] Shunsuke Yatabe. Yablo's paradox, a coinductive language and its semantics. Preprint.