Yablo's paradox, a coinductive language and its semantics

Shunsuke Yatabe*

Graduate School of Letters, Kyoto University, Japan
shunsuke.yatabe@gmail.com

Abstract. We generalize the framework of Barwise and Etchmendy's "the liar" to that of coinductive language, and focus on two problems, the mutual identity of Yablo propositions coded by hypersets in **ZFA** and the difficulty of constructing semantics. We define a coding as a game theoretic syntax and semantics, which can be regarded as a version of Austin semantics.

1 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. The answer seems to depend on how these propositions are constructed, and the essence of their construction. Yablo propositions satisfy a characteristic property in which 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,

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

This means each S_i is constructed by directly using S_{i+1} and $\neg S_{i+1}$. However, to construct S_{i+1}, we need S_{i+2} and $\neg\mathbf{Tr}(\ulcorner S_{i+2} \urcorner)$, 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 (1) we only directly use finitely many already-constructed objects to construct a new object, (2) we need infinitely many steps to reach the *initial* construction case (this is not inductive construction). Such constructions are called *coinductive*, and are widely used

* This research was partially supported by the Japanese Ministry of Education, Culture, Sports, Science and Technology, Grant-in-Aid for Scientific Research (C) 23520008, 2012.

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 an evidence that coinduction is naturally used in natural language. There are some theories which allow such construction. One of the most famous theories is **ZFA**. It is **ZF** minus the axiom of foundation plus the axiom of Anti-Foundation (**AFA**) which allows to define *hypersets*, non-wellfounded sets. We can code Yablo propositions by hypersets in **ZFA** easily.

The languages which allow to construct formulae coinductively, i.e. are to have sentences of infinite length, are called *coinductive languages*. Recently some philosophers begun to study such languages: Leitgeb studied an instance whose propositions are coded by hypersets in **ZFA** [L04]. Contrary to the inductive languages, the aspects of coinductive languages are not still well-known, in particular with respect to identity. For example, the standard identity relation over coinductive objects is *bisimulation* which is an observational equivalence of behaviors of automatons. However, in the context of Yablo's paradox, it involves some unexpected results. For, as Yablo pointed out in [Yab06], there is a counter-intuitive problem that any propositions S_i, S_j of Yablo's paradox are mutually identical. Furthermore, there are no suitable semantics for coinductive propositions since traditional semantics is designed for ordinary inductively-defined propositions (for example the semantics which is investigated in [L01] is contradictory).

In this paper, we focus on these problems, the mutual identity of Yablo propositions coded by hypersets in **ZFA** and the difficulty of semantics. We analyze the first problem from a game theoretic viewpoint, and solve this by taking account of *situations*, history of the game in progress: we will define a game theoretic semantics, which can be regarded as a version of Austin-like semantics in [BE87]. Every Yablo propositions are pairwise distinct there. Simultaneously, we define propositions as in Austin semantics in which Yablo propositions are not contradictory. Our result provides a framework of analyzing Yablo propositions and other coinductive propositions in a standard way, and would contribute investigations of coinductive languages.

The structure of this paper is as follows. We introduce **ZFA** in section 2.1 and a simple coding way of coinductive propositions by hypersets in section 2.2. We also introduce the problem of the mutual identity of Yablo propositions (in Russellian style) there. Next we focus on the problem, and analyze it introducing a game theoretic interpretation in section 3. We solve the problem introducing situations and Austin-like types in section 4.1. We define a game theoretic interpretation in section 4.2. Also we give a semantics, a version of Austin semantics in [BE87], in which Yablo propositions are not contradictory in section 4.3. Last, as a related topic, we give a consideration what kind of effects our result gives to the problem whether Yablo's paradox is self-referential or not in section 5.

2 Preliminaries

2.1 ZFA and "the liar"

One of the most famous ways to define a coinductive language is to use **ZFA** [BE87] [BM96]. This is done by coding coinductively 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**), and allows to define *hypersets*, which need not be well founded in classical logic. Actually **ZFA** is a set theory whose sets are constructed by co-induction in some transfinite induction step [V04]. The universe of **ZFA** is constructed by

- $\mathbf{V}_0 = \emptyset$,
- $\mathbf{V}_{\alpha+1} = \mathbf{V}_\alpha \bigcup \mathcal{P}^*(\mathbf{V}_\alpha)$,
- $\mathbf{V}_\gamma = \bigcup_{\delta<\gamma} \mathbf{V}_\delta$ for any γ limit,

where $\mathcal{P}^*(A) = \{x :\in |_{\mathbf{TC}(x)}$ is bisimilar to R for some $R \subseteq \mathbf{TC}(A)^2\}$.

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. We note that the set of urelements are denoted by U.

Theorem 1. ZFA *guarantees that any flat system of equations defines hypersets uniquely.*

As a sort of coinductive definition, 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 coinductive construction of formulae [1]. 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

[1] Many theories allow coinductive 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 coinductive characters [Yat12]. However, **ZFA** is the most well known among them.

such paradoxical propositions. The real value of this framework is that it allows many kinds of coinductive construction.

Last we introduce the identity relation over hypersets. For ordinary sets, the axiom of extensionality guarantees that two sets are identical if they share all elements. However, this is for inductively-constructed sets essentially, and this does not guarantee the uniqueness of hypersets: it cannot guarantee the $a = b$ such that $a = \{b\}$ and $b = \{a\}$ for example. So we extend the criteria of the identity to hypersets.

Definition 2 (bisimulation). *A bisimulation relation on sets is a binary relation R such that, for any a, b, if aRb then*

- *for any $c \in a$ there exists $d \in b$ such that cRd,*
- *for any $d \in b$ there exists $c \in a$ such that cRd,*
- $a \cap \mathcal{U} = b \cap \mathcal{U}$.

a and b are bisimilar if there is a bisimulation R such that aRb.

(1) says $\in$-relation in a is simulated by that in b, and (2) says vice versa. Usually it is interpreted to represent the observational equality of behaviors of automata: sets are regarded as state transition systems whose state transitions are represented by $\in$-relation. So hypersets are regarded as non-terminate automata of infinite size. Then **ZFA** proves the following theorem:

Theorem 2 (Strong extensionality). *The identity relation $=$ is the largest bisimulation relation, i.e.*

- *$=$ is a bisimulation relation,*
- *for any bisimulation relation on sets R and for any set a, b, aRb implies $a = b$.*

This means two hypersets are equal when their behaviors, as automata, are observationally equal though they are defined in different ways. This guarantees $a = b$ where $a = \{b\}$ and $b = \{a\}$ for example.

2.2 Coding Yablo propositions by hypersets

Let us introduce the construction of *Russellian propositions* or *Austinian types.*

Definition 3. *The coinductive language $\mathcal{L}_\in$ is*

- *propositional variables $x_0, x_1, \cdots$,*
- *logical connectives: $\wedge$, $\vee$, and infinite connectives $\wedge_{i \in I}$ and $\vee_{i \in I}$,*
- *negation: $\neg$,*
- *truth predicate* **Tr**.

For simplicity, we omit the quotation operator for truth predicate. Define their *coinductive* coding method by hypersets as follows:

Definition 4 (Russellian propositions or Austinian types). *Formulae are coinductively coded in* **ZFA** *as follows:*

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

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.

Example 1.
- the truth-teller proposition τ is coded by a Russellian proposition $\lceil \tau \rceil$ satisfying $x = \{\mathbf{t}, x\}$,
- the liar proposition λ is coded by a Russellian proposition $\lceil \lambda \rceil$ satisfying $x = \{\mathbf{n}, \mathbf{t}, x\}$.

Next let us define Yablo propositions. Literally to say, Yablo (Russellian) propositions $\{Y_n : n \in \omega\}$ 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 $Y_0, Y_1, \cdots$ are solutions of $x_0, x_1, \cdots$.

Let us evaluate this coding: we find two following problems. First, the analysis based on this coding does not go beyond $\neg Y_n$, i.e. we cannot analyze the form of this proposition any more. Second, to begin with, truth predicate seems not to be necessary for simulating *the structure of derivations* in Yablo's paradox in **ZFA**: a machinery of coinduction is enough to define Yablo's infinite regressive propositions in in our coinductive language. Therefore we try another coding for the simplicity: we define *positive* and *negative* propositions separately.

Definition 5 (Yablo propositions). *Yablo (Russellian) propositions* $\langle S_n : n \in \omega \rangle$ *are defined by the following equation: let* $\langle \{x_n, y_n, p_n, q_n : n \in \omega\}, \{\mathbf{c}, \mathbf{d}\}, e \rangle$ *be an infinite flat system such that, for any* $n \in \omega$*,*

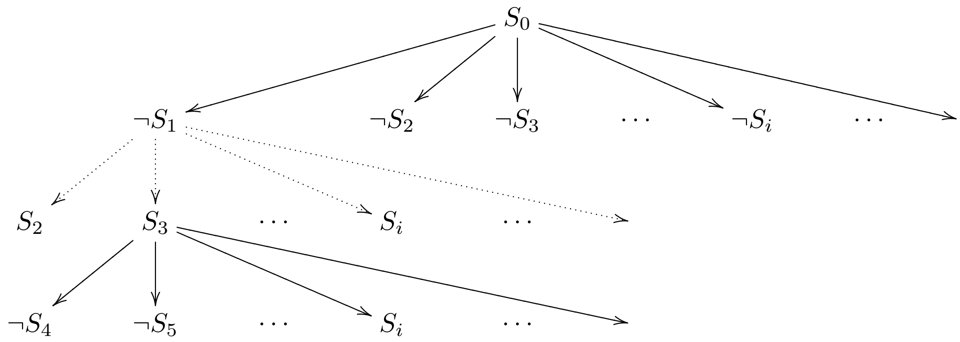
$$\begin{aligned} e(p_n) &= \{\mathbf{c}, y_n\} \\ e(q_n) &= \{\mathbf{c}, x_n\} \\ e(x_n) &= \{p_k : k > n\} \\ e(y_n) &= \{q_k : k > n\} \end{aligned}$$

$S_0, S_1, \cdots$ *are positive propositions, solutions of variables* $x_0, x_1, \cdots$*, and* $\neg S_0, \neg S_1, \cdots$ *are negative propositions, solutions of variables* $y_0, y_1, \cdots$.

The intuitive meaning of S_n is $\wedge_{n<i} \neg S_i$, and this is equivalent to $\neg S_{i+1} \wedge S_{i+1}$. Recall that the liar paradox is not unique but an instance of a self-referential paradox (the truth predicate is just a machinery to realize the self-referentiality); a Russell paradox is another (the comprehension principle plays the role). In this sense, Yablo's paradox is just an instance of a *coinductive* paradox.

Theorem 3. *Yablo (Russellian) propositions $\langle S_n : n \in \omega \rangle$ exists in* **ZFA**.

The proof is a simple application of theorem 1. The Yablo propositions form the following infinite tree.



However, an unexpected result holds because of the strong extensionality [Yab06].

Theorem 4. *All S_n are pairwise identical, i.e. $S_i = S_j$ for any $i, j \in \omega$.*

Proof. Let us fix S_0 and S_1 for example. Let us define a bisimulation relation $\sim$ by $a \sim b \equiv (\forall x \in \omega)[a = S_x \wedge b = S_{x+1}]$. Then, for any $a \in S_0$, if $a = S_k$ for any $k > 0$, then we can find $S_{k+1} \in S_1$ (and $S_k \sim S_{k+1}$). $\square$

3 A game theoretic interpretation of Yablo propositions

As we wrote, Yablo pointed out that there is a counterintuitive problem that any propositions S_i, S_j of Yablo's paradox are mutually identical because of the strong extensionality of **ZFA** in [Yab06]. We focus on this problem in this section. Our framework here is not **ZFA** but a system which can distinguish every Yablo propositions, and we define a game theoretic interpretation of Yablo propositions there. Then we investigate a criteria of the identity of propositions which simulates the strong extensionality of **ZFA** over the interpretation.

We fix **ZF** as the framework in this section, and encode a proposition of $\mathcal{L}_\in$ as an infinite tree whose nodes are labelled by the name of the proposition, and define a game theoretic interpretation of $\mathcal{L}_\in$. Since $\mathcal{L}_\in$ has infinitary connectives, the game is more complex than standard game theoretic semantics of classical propositional logic.

Definition 6 (Game G_P). *Let P be a code of a proposition. G_P is a game played by two player I (verifier), II (falsifier) such that*

(0) **the initial step:** *I asserts P true,*

(n+1) **the successor step:** *either*

- *assume I has asserted $\langle P_j : j < m \rangle$ are true, then II can attack any of I's claim: II picks up P_j and*
  - *if P_j is of the form $\wedge_{i \in I} A_i$, II tries to refute one of them: II picks up A_i ($i \in I$) and I insists that A_i is true,*

- ∗ *if P_j is of the form $\vee_{i\in I}A_i$, I tries to verify it: I picks up A_i ($i \in I$) and I insists that A_i is true,*
- ∗ *if P_j is of the form $\neg A$, I picks up A and I insists that A is false,*
- ∗ *if P_j is of the form $\mathbf{Tr}(\ulcorner A \urcorner)$, I picks up A and I insists that A is true,*

- *assume I has asserted $\langle P_j : j < m\rangle$ are false,*
  - ∗ *if P_j is of the form $\wedge_{i\in I}A_i$ or $\vee_{i\in I}A_i$, just do the same play with the case I asserts $\vee_{i\in I}\neg A_i$ or $\wedge_{i\in I}\neg A_i$ are true by using infinite de Morgan law,*
  - ∗ *if P_j is of the form $\neg A$, I picks up A and I insists that A is true,*
  - ∗ *if P_j is of the form $\mathbf{Tr}(\ulcorner A \urcorner)$, I picks up A and I insists that A is false,*

In any case, II does not have to assert something in every steps.

(ω) *II wins if I's assertion is contradictory: I picks up some proposition P" and asserts it is true (false) which has already been asserted by I to be false (true). Otherwise I wins.*

Example 2 (Game G_{S_n}). Let S_n be a code of an Yablo proposition S_0. G_{S_n} is a game played by two players I, II such that

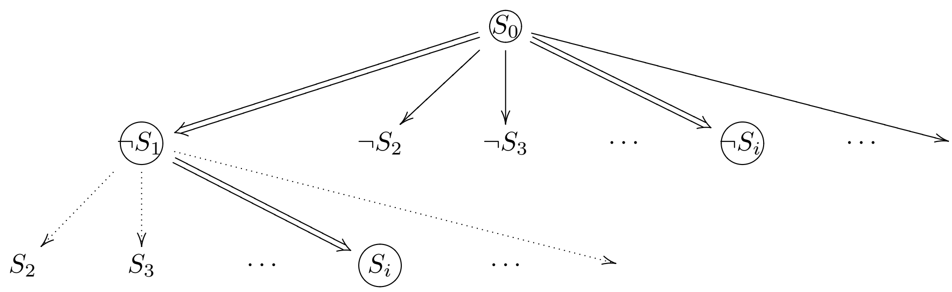
(0) I asserts S_n is true where $S_n \equiv \wedge_{i>n}\neg S_i$,

(1) the first challenge: II attacks I's assertion of "S_n is true",
- II picks up $j_0 > n$ and asserts that $\neg S_{j_0}$ is not true where $\neg S_{j_0} \equiv \vee_{k>n}S_k$,
- I picks up $k_0 \geq j_0$ and asserts that S_{k_0} is true

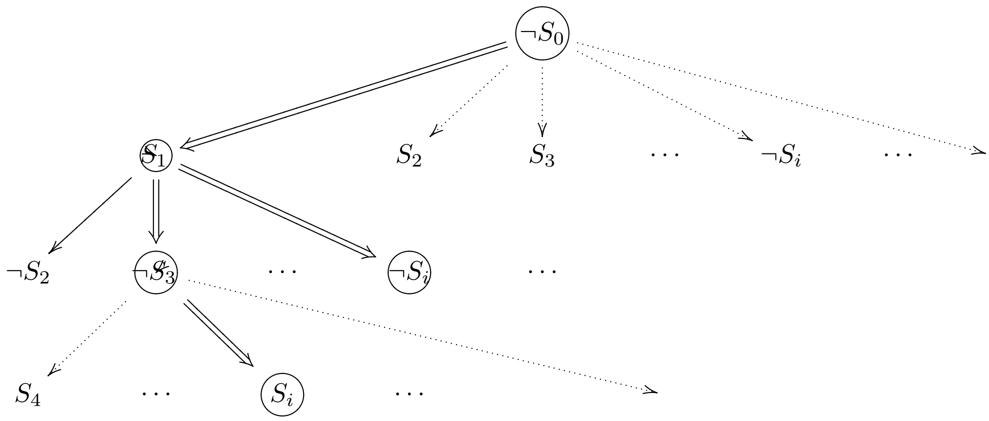
(2) the second challenge: II has two choices: (s)he can attack either "S_n is true" or "S_{k_0} is true". In the latter case,
- II picks up $j_1 \geq k_0$ and asserts that $\neg S_{j_1}$ is not true,
- I picks up $k_1 \geq j_1$ and asserts that S_{k_1} is true,

(ω) II wins if II picks up j_1 at step m asserted by I to be true; otherwise I wins.

The basic idea of G_{S_n} is this: II can challenge any branch of conjunction, and I has to refute II's claim by showing one branch of disjunction is true. At the same time I should be consistent with his/her history. The following picture shows a play history of G_S (continuous lines represent conjunction and dash lines represent disjunction):



Similarly, the following picture shows a play history of $G_{\neg S}$:



Yablo's paradox just says the following:

Theorem 5. *II has a winning strategy for both* G_{S_0} *and* $G_{\neg S_0}$.

Proof. The winning strategy for II is very easy in G_{S_0}: fix j_1 such that $j_1 = k_0$ at the second challenge. Let us give an example: if first I chooses $\neg S_1$ at the second level at (1), and next if I chooses S_i then II chooses $\neg S_i$ at the first level at (2) (see the following picture). The case of $G_{\neg S}$ is similar. □

The fact we regard every Yablo propositions mutually identical corresponds to focusing that all games on Yablo propositions are isomorphic and ignoring the difference of the respective moves of the players. In G_S, first I chooses $n \in \omega$. Whatever (s)he does, (s)he loses the game after two steps, and the both histories of these games are observationally equal *except the choice of index.* This is guaranteed by the fact that all Yablo propositions are mutually bisimilar, therefore the first choice of $n \in \omega$ does not give any effect on victory or defeat of the game: this means whatever (s)he first choose, the *behavior* of the paradox, the derivation of the inconsistency, is of the identical form. In this sense the equality of the play of the game implies the equality of propositions itself in **ZFA**.

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 $\neg S \to S$. In this sense, Yablo's paradox is a sort of a self-referential paradox (we will discuss the detail of this in section 5).

4 An Austin-like semantics with situations

As we saw, any Yablo propositions are mutually identical. For some philosophers like Yablo, this is a counterintuitive consequence because it seems to be intuitive to think that all Yablo propositions are pairwise distinct. Let us remember the similar problem happens when Russellian semantics had been developed [BM96]: they introduced Austin semantics to avoid this problem at the same time. Roughly speaking, an Austinian proposition is a pair of a *situation* and an *Austinian type*. Austinian types are just Russellian propositions. Situations make intuitively different propositions really different, i.e. for any type τ and situations s, s', $\langle s, \tau\rangle$ and $\langle s', \tau\rangle$ are different propositions. Furthermore, Austinian semantics is consistent, the liar proposition is just false, though Russellian semantics is contradictory by the liar paradox.

This approach can be applied to our coinductive language: the variation of Austin semantics can be used to avoid the counter-intuitivity. Basic ideas are as follows:

- Austin-like types are trees whose nodes are indexed by natural numbers. They are to simulate games like G_{S_n} in section 3 by games over them introduced in this section,
- Situations represent the different plays, or histories, of the game. Yablo propositions.

We introduce Austin-like types and Yablo trees $\mathbf{X}_0, \mathbf{Y}_0, \cdots$ (which is the austin-like type of the Yablo proposition) in section 4.1, next we define the game over Yablo tree which is an analogue of game G_{S_0} in section 4.2, last we define propositions and semantics in which Yablo propositions are not contradictory but false in section 4.3.

4.1 Austin-like types

First we introduce a notation on *Austin-like types*. We have already shown that, contrary to our intentions, our coding of $\mathcal{L}_\in$ collapse all distinct Yablo propositions to be the same proposition S. This is a technical problem of our coding, and we can resolve this by fixing a new coding method. Therefore we define a new interpretation way and the *Yablo trees* $\mathbf{X}_0, \mathbf{Y}_0, \cdots$ indexed by natural numbers to simulate the game G_S in example 2 by new game $\mathbf{G}_{\mathbf{X}_0}$.

Let us introduce a new interpretation: we define the basic idea of the interpretation from formulae of $\mathcal{L}_\in$ to hypersets by set equations as follows:

- First we define two sorts of variables (urelements) in addition to standard variables:
  - positive variables: $x_0, x_1, \cdots$,
  - negative variables: $y_0, y_1, \cdots$

  If P is interpreted to a positive variable x_j, then $\neg P$ is interpreted to the negative variable y_j (and vice versa).
- To encode the fact that P is interpreted to x_j (y_j), x_j is defined to include *indexes* $\langle 1, j\rangle$ ($\langle 0, j\rangle$); 1 (0) is a sign of the positivity (negativity), and j represents it is the nth variable.

Let us give an example.

Example 3. Let us interpret the liar proposition λ. Let λ be positive: let us express it x_0. Then $\neg\lambda$ is negative (represented by y_0). Let $\langle\{x_0, y_0, p_0, q_0\}, \{\langle 0,0\rangle, \langle 1,0\rangle\}, e\rangle$ be an infinite flat system such that, for any $n \in \omega$,

$$\begin{aligned} e(p_0) &= \{y_0\} \\ e(q_0) &= \{x_0\} \\ e(x_0) &= \{p_0, \langle 0,0\rangle\} \\ e(y_0) &= \{q_0, \langle 1,0\rangle\} \end{aligned}$$

Here p_0, q_0 are standard variables. Here, the set which satisfies x_0 is the code of the liar proposition. The set which satisfies y_0 is the code of the negation of the liar proposition.

Definition 7 (Yablo trees). *The Yablo trees* $\mathbf{X}_0, \mathbf{X}_1, \cdots \mathbf{Y}_0, \mathbf{Y}_1, \cdots$ *are defined as follows: let* $\langle\{x_n, y_n, p_n, q_n : n \in \omega\}, \{\mathbf{c}, \mathbf{d}\}, e\rangle$ *be an infinite flat system such that, for any* $n \in \omega$,

$$\begin{aligned} e(p_n) &= \{\mathbf{c}, y_n\} \\ e(q_n) &= \{\mathbf{c}, x_n\} \\ e(x_n) &= \{p_k, \langle 0,n\rangle : k > n\} \\ e(y_n) &= \{q_k, \langle 1,n\rangle : k > n\} \end{aligned}$$

$\mathbf{X}_0, \mathbf{X}_1, \cdots$ *are solutions of* $x_0, x_1, \cdots$ *and represent Yablo propositions* $S_0, S_1, \cdots$, *and* $\mathbf{Y}_0, \mathbf{Y}_1, \cdots$ *are solutions of* $y_0, y_1, \cdots$ *and represent* $\neg S_0, \neg S_1, \cdots$.

We recall $\mathbf{c}$ just represents the conjunction and $\mathbf{d}$ does the disjunction. Here, $\mathbf{X}_i, \mathbf{Y}_i$ are indexed by i, i.e. $\langle i,1\rangle \in \mathbf{X}_i, \langle i,0\rangle \in \mathbf{Y}_i$ and $\langle j,1\rangle \notin \mathbf{X}_i, \langle j,0\rangle \in \mathbf{Y}_i$ if $i \neq j$. This makes every Yablo trees mutually different because there is no bisimulation between i and j if $i \neq j$. We call these interpreted formula *Austin-like Type.*

We note that any $\mathbf{X}_n$ forms an infinite-branching tree of infinite height (and so is $\mathbf{Y}_n$) such that for any i, j, $\mathbf{X}_i$ and $\mathbf{X}_j$ are isomorphic. Each tree S_n is *self-similar*, i.e., for any branch t of $\mathbf{X}_n$, there is a sub-tree $T \subseteq \mathbf{X}_n|_t$ such that there is an isomorphism $\pi_j : T \to \mathbf{X}_j$ for some $j > n$. Such self-similarity is a specific character of coinductive object: the tree and isomorphisms form a completely iterative algebra [Mo08]. It is easy to see that the self-similarity is a key of this paradox: whatever the player I choose, the plays are pairwise isomorphic because the subtrees over the nodes chosen by the player I are pairwise isomorphic.

4.2 A game theoretic interpretation

Next let us define a game theoretic interpretation of types as in the previous section. It is easy to simulate G_S by using $\mathbf{X}_0$.

Definition 8 (Game $G_{\mathbf{X}_0}$). $G_{\mathbf{X}_0}$ *is a game played by two player I (verifier), II (falsifier) such that*

(0) **the initial step:** *I asserts $\mathbf{X}_0$ true,*

(n+1) **the successor step:** *either*

- *assume I has asserted $\langle T_j : j < m\rangle$ are true, then II can attack any of I's claim: II picks up T_k and*
  - *if T_k is of the form $\{\langle \mathbf{c}, i, P_i\}$, II tries to refute one of them: II picks up P_i $(i \in \omega)$ and I insists that P_i is true,*
  - *if T_k is of the form $\{\langle \mathbf{d}, i, P_i\}$, I tries to verify it: I picks up P_i $(i \in \omega)$ and I insists that P_i is true,*

in any case, II does not have to assert something in every steps.

(ω) II wins if I's assertion is contradictory: I picks up some proposition $\mathbf{X}_i$ ($\mathbf{Y}_i$) and asserts it is true (false) which has already been asserted by I to be false (true). Otherwise I wins.

As in example 2, the essence of the paradox is that, for any node of $\mathbf{X}_0$ and any branch T of that node, if T is of the form $\mathbf{X}_j$ we can find $\mathbf{Y}_j$ in the first node (otherwise T is of the form $\mathbf{Y}_j$ and we can find $\mathbf{X}_j$ in the node of the second level). This is possible because any $\mathbf{X}_j$ is pairwise isomorphic (actually we can find embedding $X_j \to \mathbf{X}_0$ for any j). We will see that such self-similarity plays an essential role in the paradox.

Definition 9 (Situation). *A situation, or a history of $\mathbf{G}_{\mathbf{X}_0}$ is of the form $\langle\langle k, i\rangle\rangle : n \in \omega\rangle$ where $n+1, k, i$ are in definition 8.*

Definition 10. – *For any situation s, $s \models P$ if I insists P is true when the game history is s in the game $G_{\mathbf{X}_0}$,*

– *s is actual situation if s is a situation such that $s \not\models \mathbf{X}_i$ or $s \not\models \mathbf{Y}_i$ for any $i \in \omega$.*

Example 4. Let us consider the following play of the game:

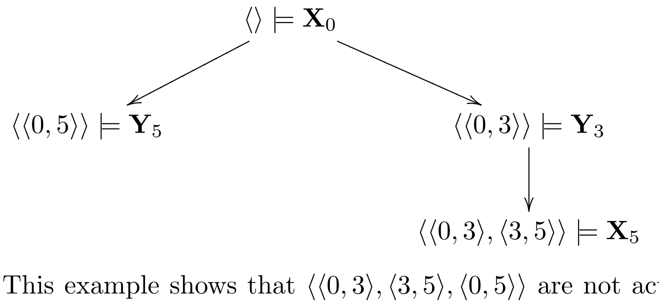
(0) I asserts $\mathbf{X}_0$ is true; the history is $\langle\rangle$,

(1) II attacks to $\mathbf{X}_0$: II picks up 3 and I claims that $\mathbf{Y}_3$ is true; the situation is $\langle\langle 0, 3\rangle\rangle$,

(2) II attacks to $\mathbf{Y}_3$: I picks up 5 and (s)he insists that $\mathbf{X}_5$ is true; the situation is $\langle\langle 0, 3\rangle, \langle 3, 5\rangle\rangle$,

(4) II attacks to $\mathbf{X}_0$ again: II picks up 5 and I claims that $\mathbf{Y}_5$ is true; the situation is $\langle\langle 0, 3\rangle, \langle 3, 5\rangle, \langle 0, 5\rangle\rangle$

(ω) I looses the game because I insists both $\mathbf{X}_5$ and $\mathbf{Y}_5$ are true.



This example shows that $\langle\langle 0, 3\rangle, \langle 3, 5\rangle, \langle 0, 5\rangle\rangle$ are not actual situation, i.e. contradictory.

4.3 Propositions and semantics

Last let us define propositions and semantics along the line of [BE87].

Definition 11 (Proposition).

- *P is a proposition if P is of the form $\langle s, T\rangle$ where s is a situation and T is an Austin-like Type,*
- *Q is a contradictory proposition if Q is of the form $\langle s, T\rangle$ such that s is not actual.*

Definition 12 (Semantics). *For any proposition P,*

- *P is true (or $\models P$) if $P = \langle s, T\rangle$ and $s \models T$,*
- *P is false otherwise.*

Example 5. $\langle\langle\langle 0, 5\rangle\rangle, \mathbf{X}_5\rangle$ is a proposition (which is not contradictory), though $\langle\langle\langle 0, 3\rangle, \langle 3, 5\rangle, \langle 0, 5\rangle\rangle, \mathbf{X}_5\rangle$ is a contradictory proposition.

In this semantics, Yablo propositions are not intrinsically contradictory: it depends on the fact that what situation we are thinking of. This is the same as the liar paradox in [BE87].

Let us summarize our new coding and semantics. In this setting,

- all Yablo types $\mathbf{X}_0, \mathbf{X}_1, \cdots, \mathbf{Y}_0, \mathbf{Y}_1, \cdots$ (which are Austin-like types) are pairwise distinct because of the indexes,
- Yablo types are not always contradictory, i.e. for any actual situation s, both $\langle s, \mathbf{X}_i\rangle$ and $\langle s, \mathbf{Y}_i\rangle$ are actual.

5 Discussion: Yablo's paradox and self-referentiality

In this section, we discuss the philosophical connotations of our coinductive construction. One of the most well-discussed issue of Yablo's paradox is whether Yablo's paradox is self-referential or not. On the one hand, actually **ZFA** provides a natural model of self-referentiality: all Yablo propositions defined in **ZFA** are mutually identical since there exists a bisimulation, a natural identity relation on behaviors of automatons, over them. On the other hand, just adding indexes to Yablo propositions makes them pairwise different

Let us remember the context of this problem. It is often said that circularity is source of the inconsistency caused by paradoxes. At least, the full form of self-reference causes a trouble, the liar paradox, "This sentence is false". Then it is natural to ask whether there is any non-self-referential paradox or not. Then Stephan Yablo insisted that the self-referentiality is not necessary for paradoxes [Yab93]. His proposal was very controversial since he never considered how we can construct these sentences. Therefore Graham Priest objected that the paradox is actually self-referential because the diagonalization is used to construct the sentences in truth theories [P97]. The controversy between him and Graham Priest, and the proxy war between Sorensen and JC Beal are well-known, but seem not to be productive: it is like "it is self-referential in this sense" and "it is not self-referential in that sense" from beginning to end. As Hannes Leitgeb pointed out, *two different notions of self-referentiality and circularity have been used* in the controversy [L01]. The real issue was on the definition of self-referentiality, and we can ask *what might a formally correct and materially adequate definition of self-referentiality look like?*. It is plausible that the definition of self-referentiality is different if the base theory is different and it should be difficult to give a formally correct definition of self-referentiality which is common to many different kinds of theories (truth theories, etc.), therefore we can ask *does every such theory have its 'own' formal concept of self-referentiality?*

Of course the problem of the formal definition of self-referentiality is the problem of the formal definition of the identity. The basic thesis of this paper is that we cannot consider the criteria of identity without thinking of how these objects are constructed. Therefore, to answer above questions, we have to find the common core concept underlying various constructions of Yablo sentences: as we have already seen, it is *coinduction.* To emphasize this clearly, we constructed Yablo propositions by purely coinductive way in **ZFA**, which allows coinductive definitions in a strong form, in section 2.2, and we constructed *Yablo tree*, a domain of the game which represents the derivation in the paradox in section 4. Yablo tree is a self-similar object, and it is a characteristic property of coinductively defined objects.

The problem whether Yablo's paradox is self-referential should be thought as a problem what is a criteria of the identity of coinductive objects. In the context of coinduction, usually *bisimilarity* is fixed as the criteria: it is interpreted as the observational equality of two automatons, i.e. the similarity of the behaviors of automatons, and it is realized by the strong extensionality in **ZFA**.

As we saw in section 3, if we naively code Yablo propositions by hypersets, then they are mutually identical since the plays (or behaviors) of games over Yablo trees are absolutely isomorphic. This result may involve a new viewpoint on a criteria of the identity of coinductive propositions, the viewpoint which regards *propositions as automata.* Here propositions are represented by automata or state transition systems, and the word "*behaviors*" means derivations of propositions from other propositions. Non-terminate automata are typical coinductive objects, and in the context of propositions, they represent propositions whose truth value calculation never terminates, i.e. *ungrounded* propositions. Coinduction is a mathematical method which gives precise description of behaviors of Kripke's ungrounded propositions in detail. Our view suggests that the identity criteria of propositions should be observational equality, i.e. two automata are observationally equal if their behaviors are equal. Let us consider the following propositions:

$$\begin{aligned} \lambda &\equiv \neg\mathbf{Tr}(\ulcorner\lambda\urcorner) \\ \lambda' &\equiv \neg\mathbf{Tr}(\ulcorner\lambda'\urcorner) \wedge \ulcorner\lambda\urcorner \neq \ulcorner\lambda'\urcorner \end{aligned}$$

Both λ and λ' can be seen as instances of the liar proposition because their behaviors are identical though they are literally different. If not, we should ask which is the "true" liar sentence. The problem of this account is that, according to Yablo, this is *counter-intuitive* [Yab06]: we may think S_1 and S_0 are different intuitively.

On the other hand, we can code Yablo propositions by pairwise distinct sets by just adding indexes as we saw in section 4. In this setting, Yablo's paradox is not self-referential: all types of Yablo propositions, Yablo trees, are *self-similar* and pairwise isomorphic (but not bisimilar because of the existence of indexes). Let us remember that the self-similarity is an essential property of coinductive objects [Mo08]. This account involves that the essence of the paradox, which does not depend on how the paradox is formalized, is self-similarity and whether Yablo's paradox is self-referential or not is less-essential.

If we really solve this self-referentiality problem, we have to suspect Yablo's intuition: the intuition, Yablo propositions are pairwise distinct, might be an illusion caused by indexes. Rather we had better to regard this as a pseudo-problem: we should concentrate on specialities of coinductive objects like self-similarity since the essence of the paradox is self-similarity, and regard that the self-referentiality is contingent, i.e. it is self-referential in the former setting though it is not so in the latter.

References

[A88] Aczel, Peter. Non-well-founded sets. CSLI publications (1988)

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

[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.

[L04] Hannes Leitgeb. Circular languages. Journal of Logic, Language and Information 13 (2004) pp.341-371.

[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.

[Yat12] Shunsuke Yatabe. A constructive naive set theory and infinity. Preprint.