Yabloesque Paradoxes and Modal Logic

Ahmad Karimi

Department of Mathematics

Behbahan Khatam Alanbia University of Technology

Behbahan-IRAN

Workshop on

Various Aspects of Modality

May 12, 2016

Overview

Yablo's Paradox

Yablo's Paradox

Yablo considers the following sequence of sentences $\{S_i\}$:

$$S_1 : \forall k > 1; \; S_k \text{ is untrue,}$$
$$S_2 : \forall k > 2; \; S_k \text{ is untrue,}$$
$$S_3 : \forall k > 3; \; S_k \text{ is untrue,}$$
$$\vdots$$

Linear Temporal Logic (LTL)

The inductive definition of formulas (of *LTL*) is as follows[1]:

$$\varphi ::= \text{false} \mid c \mid \neg\varphi \mid \varphi \to \varphi \mid \bigcirc\varphi \mid \Box\varphi$$

$\bigcirc\varphi$: "next φ" $\Box\varphi$: "always φ" $\Diamond\varphi$: "sometimes φ"

The Intended Model: Kripke model $\langle \mathbb{N}, \Vdash \rangle$ where $\Vdash \subseteq \mathbb{N} \times \texttt{Atoms}$ can be extended to all formulas by:

- $n \Vdash \varphi \wedge \psi$ iff $n \Vdash \varphi$ and $n \Vdash \psi$
- $n \Vdash \neg\varphi$ iff $n \nVdash \varphi$
- $n \Vdash \bigcirc\varphi$ iff $n+1 \Vdash \varphi$
- $n \Vdash \Box\varphi$ iff $m \Vdash \varphi$ for every $m \geq n$

Theorem

$LTL \models \varphi$ if and only if $\neg\varphi$ is not satisfiable.

[1] **F. Kröger**, **S. Merz**, *Temporal Logic and State Systems*, Springer, 2008.

Yablo's Paradox as a Theorem in LTL

Here, for the very first time, we use this paradox (actually its argument) for proving some genuine mathematical theorems in LTL.

The thought is that we can make progress by thinking of the sentences in the statement of Yablo's paradox not as an infinite family of atomic propositions but as a single proposition evaluated in lots of states in a Kripke model.

Thus the derivability of Yablo's paradox should be the same fact as the theoremhood of a particular formula in the linear temporal logic.
A version of Yablo's paradox is a sentence S that satisfies the following equivalence[2]

$$\Box(S \longleftrightarrow \bigcirc\Box\neg S)$$

[2] **A. Karimi** & **S. Salehi**, *Diagonal Arguments and Fixed Points*, Bulletin of the Iranian Mathematical Society, to appear.

Yablo's Paradox as a Theorem in LTL

Theorem (Yablo's Paradox)

LTL $\models \neg\Box(\varphi \leftrightarrow \bigcirc\Box\neg\varphi)$.

Proof. To prove the formula $\neg\Box(\varphi \leftrightarrow \bigcirc\Box\neg\varphi)$ is valid in LTL, we need to show the formula $\Box(\varphi \leftrightarrow \bigcirc\Box\neg\varphi)$ is not satisfiable.

For a moment assume that there is a Kripke model $\langle \mathbb{N}, \Vdash \rangle$ and $n \in \mathbb{N}$ for which $n \Vdash \Box(\varphi \leftrightarrow \bigcirc\Box\neg\varphi)$. Then $\forall i \geq n : \; i \Vdash (\varphi \leftrightarrow \bigcirc\Box\neg\varphi)$. We distinguish two cases:

(1) For some $j \geq n$, $j \Vdash \varphi$. Then $j+1 \Vdash (\Box\neg\varphi)$ so $j+l \nVdash \varphi$ for all $l \geq 1$. In particular $j+1 \nVdash \varphi$ whence $j+2 \nVdash (\Box\neg\varphi)$ which is in contradiction with $j+1 \Vdash (\Box\neg\varphi)$.

(2) For all $j \geq n$, $j \nVdash \varphi$. So $n \nVdash \varphi$ and $n+1 \nVdash (\Box\neg\varphi)$ hence there must exist some $i \geq n$ with $i \Vdash \varphi$ which contradicts by (1) above. Thus, the formula $\Box(\varphi \leftrightarrow \bigcirc\Box\neg\varphi)$ cannot be satisfiable in LTL.

Yablo's Paradox as a Theorem in LTL

Other Versions of Yablo's Paradox

Yablo's paradox comes in several varieties; here we show that other versions of Yablo's paradox become interesting theorems in LTL as well.

(always): $\mathcal{Y}_n \iff \forall i > n \, (\mathcal{Y}_i \text{ is not true })$

(sometimes): $\mathcal{Y}_n \iff \exists i > n \, (\mathcal{Y}_i \text{ is not true })$.

(almost always): $\mathcal{Y}_n \iff \exists i > n \, \forall j \geq i \, (\mathcal{Y}_j \text{ is not true })$.

(infinitely often): $\mathcal{Y}_n \iff \forall i > n \, \exists j \geq i \, (\mathcal{Y}_j \text{ is not true })$.

Theorem

$\mathrm{LTL} \models \neg\Box(\varphi \leftrightarrow \bigcirc\Diamond\neg\varphi)$ (Sometimes Yablo's Paradox)

$\mathrm{LTL} \models \neg\Box(\varphi \leftrightarrow \bigcirc\Diamond\Box\neg\varphi)$ (Almost Always Yablo's Paradox)

$\mathrm{LTL} \models \neg\Box(\varphi \leftrightarrow \bigcirc\Box\Diamond\neg\varphi)$ (Infinitely Often Yablo's Paradox)

Yablo's Paradox in Modal Logic KD4

Yablo's paradox can be formalized in modal logic. One can demystify Yablo's paradox by showing that it can be thought of as the fact that the formula $\Box(\varphi \leftrightarrow \Box\neg\varphi)$ is unsatisfiable in a modal logic characterized by frames that are strict partial orders without maximal elements.

Recall that Yablo's paradox is in the form:

$$(\forall i)\big[S(i) \longleftrightarrow (\forall j > i) : \neg S(j)\big] \qquad (*)$$

where the variables range over $\mathbb{N}$. The only assumptions we have used are as follows:

- $>$ is transitive
- $>$ is irreflexive
- $(\forall i)(\exists j)(j > i)$

Yablo's Paradox in Modal Logic KD4

The suitable logic is normal modal logic KD4 where K is $\Box(\varphi \to \psi) \to (\Box\varphi \to \Box\psi)$, D is $\Box\varphi \to \Diamond\varphi$ and 4 is $\Box\varphi \to \Box\Box\varphi$. The axiom D characterizes seriality $(\forall x)(\exists y)R(x,y)$ and the axiom 4 characterizes transitivity $(\forall xyz)(R(x,y) \wedge R(y,z) \to R(x,z))$.

A sequent calculus rules for the logic KD4 are presented in (Forster and Gore 2015)[3].

Theorem

The formula $\neg\Box(\varphi \leftrightarrow \Box\neg\varphi)$ is KD4-valid.

[3] **T. Forster**, **R. Gore**, *Yablo's Paradox as a Theorem in Modal Logic*, Logique et Analyse 59, (2016).

Brandenburger-Keisler Paradox

Brandenburger-Keisler Paradox

In game theory, the notion of a player's beliefs about the game (even a player's beliefs about other players' beliefs, and so on) arises naturally.

Take the basic game-theoretic question: Are Ann and Bob rational, does each believe the other to be rational, and so on?

To address this, we need to write down what Ann believes about Bob's choice of strategy to decide whether she chooses her strategy optimally given her beliefs (i.e., whether she is rational).

Brandenburger-Keisler Paradox

In 2006, Adam Brandenburger and H. Jerome Keisler discovered a Russell-style paradox. The statement of the paradox involves two concepts:

beliefs and assumptions.

An assumption is assumed to be a strongest belief.

Brandenburger-Keisler Paradox

Suppose there are two players, Ann and Bob, and consider the following description of beliefs:

Ann believes that Bob assumes that
Ann believes that Bob's assumption is wrong.

A paradox arises when one asks the question[4]:

Does Ann believe that Bob's assumption is wrong?

[4] **A. Brandenburger**, **H. Jerome Keisler**, *An Impossibility Theorem on Beliefs in Games*, Studia Logica 84, (2006), 211–240.

Brandenburger-Keisler Paradox

Suppose that answer to the above question is "yes". Then according to Ann, Bob's assumption is wrong. But, according to Ann, Bob's assumption is Ann believes that Bob's assumption is wrong. However, since the answer to the above question is "yes", Ann believes that this assumption is correct. So, Ann does not believe that Bob's assumption is wrong.

Therefore, the answer to the above question must be "no". Thus, it is not the case that Ann believes that Bob's assumption is wrong. Hence Ann believes Bob's assumption is correct. That is, it is correct that Ann believes that Bob's assumption is wrong. So, the answer must have been yes. This is a contradiction.

Just as Russell's paradox suggests that not every collection can constitute a set, the Brandenburger-Keisler paradox suggests that not every description of beliefs can be "represented".

Brandenburger-Keisler Paradox in Modal Logic

BK Paradox in Modal Logic

Two-player Brandenburger-Keisler paradox can be reformulated to an impossibility result in a modal logic setting. For each pair of players cd among Ann and Bob, there will be an operator $\mathcal{B}^{cd}$ of beliefs for c about d, and an operator $\mathcal{A}^{cd}$ of assumptions for c about d.

Definition

An **interactive frame** is a structure $\mathcal{W} = (W, P, U^a, U^b)$ with a binary relation $P \subseteq W \times W$ and disjoint sets U^a, U^b, such that $\mathcal{M} = (U^a, U^b, P^a, P^b)$ is a belief model, where $U^a \cup U^b = W,\ P^a = P \cap U^a \times U^b$, and $P^b = P \cap U^b \times U^a$.

BK Paradox in Modal Logic

Interactive modal logic will have two distinguished proposition symbols $\mathbf{U}^a, \mathbf{U}^b$ and a set L of additional proposition symbols. By a modal formula we mean an expression which is built from proposition symbols and the false formula $\perp$ using propositional connectives, the universal modal operator A, and the modal operators $\mathcal{B}^{cd}$, $\mathcal{A}^{cd}$ where c and d are taken from $\{a, b\}$. Given a valuation V on $\mathcal{W}$, the notion of a world w being true at a formula φ ($w \models \varphi$), is defined by induction on the complexity of φ as follows: $w \models \mathbf{U}^a$ if $w \in U^a$, and similarly for b. That is, $\mathbf{U}^a$ is true at each state for Ann, and $\mathbf{U}^b$ is true at each state for Bob.

BK Paradox in Modal Logic

The rules for connectives are as usual, and the rules for the modal operators for each pair of players $c, d \in \{a, b\}$ are:

- $w \models \mathcal{B}^{cd}\varphi$ if $(w \models \mathbf{U}^c \wedge \forall z\,[(P(w,z) \wedge z \models \mathbf{U}^d) \longrightarrow z \models \varphi])$.
- $w \models \mathcal{A}^{cd}\varphi$ if $(w \models \mathbf{U}^c \wedge \forall z\,[(P(w,z) \wedge z \models \mathbf{U}^d) \longleftrightarrow z \models \varphi])$.

Validity and satisfiability are defined as before. If x has sort U^a and y has sort U^b, then

$$x \models \mathcal{B}^{ab}\varphi \;\text{ says }\; \text{“}x \text{ believes } \varphi(y)\text{”},$$

$$x \models \mathcal{A}^{ab}\varphi \;\text{ says }\; \text{“}x \text{ assumes } \varphi(y)\text{”}.$$

BK Paradox in Modal Logic

Definition

An interactive frame $\mathcal{W}$ with valuation V has a **hole** at a formula φ if either $\mathbf{U}^b \wedge \varphi$ is satisfiable but $\mathcal{A}^{ab}\varphi$ is not, or $\mathbf{U}^a \wedge \varphi$ is satisfiable but $\mathcal{A}^{ba}\varphi$ is not. A **big hole** is defined similarly but with $\mathcal{B}$ instead of $\mathcal{A}$.

Frame $\mathcal{W}$ is **complete** for a set L of modal formulas if it does not have a hole in L.

For the remainder, we will always suppose that $\mathcal{W}$ is an interactive frame, **D** is a proposition symbol (for diagonal), and V is a valuation in $\mathcal{W}$ such that $V(\mathbf{D})$ is the set

$$D = \big\{w \in W : (\forall z \in W)[P(w,z) \rightarrow \neg P(z,w)]\big\}.$$

BK Paradox in Modal Logic

Proposition

If $\mathcal{A}^{ab}\mathbf{U}^{b}$ *is satisfiable then*

$$[\mathcal{B}^{ab}\mathcal{A}^{ba}\mathbf{U}^{a}] \rightarrow \mathbf{D}$$

is valid.

Proposition

$\neg\mathcal{B}^{ab}\mathcal{A}^{ba}(\mathbf{U}^{a} \wedge \mathbf{D})$ *is valid.*

Thus there is no complete interactive frame for the set of all modal formulas built from $\mathbf{U}^{a}, \mathbf{U}^{b}, \mathbf{D}$.

Yablo-like Brandenburger-Keisler Paradox

In this section, we present a non-self-referential version of Brandenburger-Keisler paradox using Yablo's reasoning. Let us consider two infinite sequence of players $\{A_i\}$ and $\{B_i\}$, and following description of beliefs:

$$\begin{array}{cc} A_1 & B_1 \\ A_2 & B_2 \\ A_3 & B_3 \\ \vdots & \vdots \end{array}$$

For all i, A_i believes that B_i assumes that
for all $j > i$, A_j believes that B_j's assumption is wrong.

Yablo-like Brandenburger-Keisler Paradox

A paradox arises when one asks the question "Does A_1 believe that B_1's assumption is wrong?"

Suppose that the answer to the above question is "no". Thus, it is not the case that A_1 believes that B_1's assumption is wrong. Hence A_1 believes B_1's assumption is correct. That is, it is correct that for all $j > 1$, A_j believes that B_j's assumption is wrong. Specially, A_2 believes that B_2's assumption is wrong.

On the other hand, since for all $j > 2$, A_j believes that B_j's assumption is wrong, one can conclude that A_2 believes B_2's assumption is correct. Therefore, in the same time A_2 believes that B_2's assumption both correct and wrong. This is a contradiction!

Yablo-like Brandenburger-Keisler Paradox

If the answer to the above question is "yes". Then according to A_1, B_1's assumption is wrong. But, according to A_1, B_1's assumption is for all $j > 1$, A_j believes that B_j's assumption is wrong.

Thus, there is $k > 1$ for which A_k believes that B_k's assumption is correct. Now we can apply the same reasoning we used before about A_k and B_k to reach the contradiction! Hence the paradox.

This paradox is a non-self-referential multi-agent version of the Brandenburger-Keisler paradox[5].

[5] **A. Karimi**, *A Non-Self-Referential Paradox in Epistemic Game Theory*, arXiv:1601.06661

Interactive Temporal Assumption Logic

Interactive Temporal Assumption Logic

In this section, we introduce an Interactive Temporal Assumption Logic (iTAL) to present an appropriate formulation of the non-self-referential Yablo-like BK paradox.

The *interactive temporal assumption language* $\mathcal{L}_{iTAL}$ contains individual propositional symbols, the propositional connectives, the linear-time operators $\bigcirc, \Box, \Diamond$ and the epistemic operators "Believe" and "Assumption": for each pair of players ij among Ann and Bob, the operator $\mathcal{B}^{ij}$ will be beliefs for player i about j, and $\mathcal{A}^{ij}$ is the assumption for i about j.

In words, $\mathcal{B}^{ij}\phi$ means that the agent i believes ϕ about j, and $\mathcal{A}^{ij}\phi$ is that the agent i assumes ϕ about agent j. The temporal operators $\bigcirc, \Box$, and $\Diamond$ are called *next time*, *always (or henceforth)*, and *sometime (or eventuality)* operators, respectively. Formulas $\bigcirc\varphi$, $\Box\varphi$, and $\Diamond\varphi$ are typically read "next φ", "always φ", and "sometime φ". We note that $\Diamond\varphi \equiv \neg\Box\neg\varphi$.

Interactive Temporal Assumption Logic

Definition

Formulas in $\mathcal{L}_{iTAL}$ are defined as follows:

$$\phi \ := p \mid \neg\phi \mid \phi \wedge \psi \mid \bigcirc\phi \mid \Box\phi \mid \mathcal{B}^{ij}\phi \mid \mathcal{A}^{ij}\phi$$

For semantical interpretations we introduce an appropriate class of Kripke models. An iTAL-Model is a Kripke structure

$$\mathcal{W} = (W, \mathbb{N}, \{P_n : n \in \mathbb{N}\}, U^a, U^b, V),$$

where W is a nonempty set, for each $n \in \mathbb{N}$, P_n is a binary relation $P_n \subseteq W \times W$ and U^a, U^b are disjoint sets such that (U^a, U^b, P^a_n, P^b_n) is a belief model, where $U^a \cup U^b = W$, $P^a_n = P_n \cap U^a \times U^b$, and $P^b_n = P_n \cap U^b \times U^a$. V : **Prop** $\rightarrow 2^{\mathbb{N}\times W}$ is a function mapping to each propositional letter p the subset $V(p)$ of Cartesian product $\mathbb{N} \times W$.

Interactive Temporal Assumption Logic

Indeed, $V(p)$ is the set of pairs (n, w) such that p is true in the world w at the moment n.

The satisfiability of a formula $\varphi \in \mathcal{L}_{iTAL}$ in a model $\mathcal{W}$, at a moment of time $n \in \mathbb{N}$ in a world $w \in W$, denoted by $\mathcal{W}_n^w \Vdash \varphi$ (in short; $(n, w) \Vdash \varphi$), is defined inductively as follows:

- $(n, w) \Vdash p \iff (n, w) \in V(p)$ for $p \in$ **Prop**,
- $(n, w) \Vdash \neg\varphi \iff (n, w) \nVdash \varphi$,
- $(n, w) \Vdash \varphi \wedge \psi \iff (n, w) \Vdash \varphi$ and $(n, w) \Vdash \psi$,
- $(n, w) \Vdash \bigcirc\varphi \iff (n+1, w) \Vdash \varphi$,
- $(n, w) \Vdash \Box\varphi \iff \forall m \geq n \ \ (m, w) \Vdash \varphi$,
- $(n, w) \Vdash \mathcal{B}^{ij}\varphi \Leftrightarrow (n, w) \Vdash \mathbf{U}^i \wedge \forall z[(P_n(w, z) \wedge (n, z) \Vdash \mathbf{U}^j) \rightarrow (n, z) \Vdash \varphi])$,
- $(n, w) \Vdash \mathcal{A}^{ij}\varphi \Leftrightarrow (n, w) \Vdash \mathbf{U}^i \wedge \forall z[(P_n(w, z) \wedge (n, z) \Vdash \mathbf{U}^j) \leftrightarrow (n, z) \Vdash \varphi])$.

Interactive Temporal Assumption Logic

Intuitively, $(n, x) \Vdash \mathcal{B}^{ab}\varphi$ says that "in time n, x believes $\varphi(y)$", and $(n, x) \Vdash \mathcal{A}^{ab}\varphi$ says that "in time n, x assumes $\varphi(y)$".

A formula is *valid* for V in $\mathcal{W}$ if it is true at all $w \in W$, and *satisfiable* for V in $\mathcal{W}$ if it is true at some $w \in W$.

In an interactive assumption model $\mathcal{W}$, we will always suppose that **D** is a propositional symbol, and V is a valuation in $\mathcal{W}$ such that $V(\mathbf{D})$ is the set

$$D = \big\{(n, x) \in \mathbb{N} \times W : \ (\forall y \in W)[P_n(x, y) \to \neg P_n(y, x)]\big\}.$$

Interactive Temporal Assumption Logic

We present our formulation of the non-self-referential Yablo-like BK paradox in the interactive temporal assumption setting.

The thought is that we can make progress by thinking of the sequences of agents in Yablo-like BK paradox not as infinite families of agents but as a two individual agents that their belief and assumption can be evaluated in lots of times (temporal states) in an temporal assumption model.

Thus, the emergence of the Yablo-like BK paradox should be the same as the derivability of a particular formula in the temporal assumption logic.

Let us have a closer look at the Yablo-like BK paradox:

For all i, A_i believes that B_i assumes that
for all $j > i$, A_j believes that B_j's assumption is wrong.

Now suppose that there are ONLY two players: Ann and Bob. Assume that A_i and B_i are the counterparts of Ann and Bob in the i^{th} temporal state. Then infinitely many statements in the Yablo-like BK paradox can be represented in just one single formula using temporal tools:

$\Box\big[$Ann believes that Bob assumes that
$\big(\bigcirc\Box$(Ann believes that Bob's assumption is wrong)$\big)\big]$

Interactive Temporal Assumption Logic

The interpretation of above formula in English can be seen as: "Always it is the case that Ann believes that Bob assumes from the next time henceforth that Ann believes that Bob's assumption is wrong".

Theorem

In an interactive temporal assumption model $\mathcal{W}$, if $\Box(A^{ab}\mathbf{U}^{b})$ is satisfiable, then

$$\Box\left[B^{ab}A^{ba}(\bigcirc\Box\mathbf{D})\right] \longrightarrow \Box\mathbf{D}$$

is valid in $\mathcal{W}$.

Theorem

In any interactive temporal assumption model $\mathcal{W}$, the formula

$$\neg\Box\left[B^{ab}A^{ba}(\mathbf{U}^{a} \wedge \bigcirc\Box\mathbf{D})\right]$$

is valid.

Thank you for your attention!