JEFFREY KETLAND

YABLO'S PARADOX AND ω-INCONSISTENCY

ABSTRACT. It is argued that Yablo's Paradox is not strictly paradoxical, but rather 'ω-paradoxical'. Under a natural formalization, the list of Yablo sentences may be constructed using a diagonalization argument and can be shown to be ω-inconsistent, but nonetheless consistent. The derivation of an inconsistency requires a *uniform* fixed-point construction. Moreover, the truth-theoretic disquotational principle required is also uniform, rather than the local disquotational T-scheme. The theory with the local disquotation T-scheme applied to individual sentences from the Yablo list is also consistent.

1. INTRODUCTION

Yablo's Paradox (Yablo 1993) may be presented as involving a denumerable sequence Y_i of sentences ('Yablo sentences'), with the following truth conditions:

(Y_0) For all $n > 0$, Y_n is not true.
(Y_1) For all $n > 1$, Y_n is not true
etc.

Is this infinite list of sentences paradoxical? Roughly speaking, a set of sentences is paradoxical if there is no way of assigning truth values to the individual sentences. More exactly, in the case of Yablo's paradox, we expect that set of associated biconditionals, of the form '$Y_n \leftrightarrow$ for all $m > n$, Y_m is not true', to be inconsistent with certain basic truth-theoretic principles, particularly disquotation.

Analogously, for the usual (strengthened) Liar sentence,

(λ) λ is not true

there is an associated biconditional, symbolized $\lambda \leftrightarrow \neg T(\ulcorner \lambda \urcorner)$, which is inconsistent with the disquotational T-sentence $T(\ulcorner \lambda \urcorner) \leftrightarrow \lambda$ for λ.[1]

Matters are nothing like as straightforward in the case of Yablo's paradox.[2] We can show that the *list* of Yablo biconditionals is not inconsistent with the relevant disquotation principles (i.e., disquotation restricted to Yablo sentences). It is, however, ω-inconsistent.

Synthese (2005) 145: 295–302

One way of obtaining an inconsistency in relation to the Yablo Paradox is as follows. First, since the sentences on the right hand side are the 'truth conditions' for the sentences named in the list, this suggests that we may express the truth conditions of the Yablo sentences *uniformly*, as a *single universally quantified proposition*:

(A) The Uniform Homogeneous Yablo Principle[3]

For all n, Y_n is true if and only if, for all $m > n$, Y_m is not true.

In a sense, this is rather like writing out the truth condition for λ in the form 'λ is true if and only if λ is not true', which is transparently inconsistent.

The Uniform Homogeneous Yablo Principle (A) is inconsistent. For suppose that Y_0 is true. Then, by (A), for all $m > 0$, Y_m is not true (*). *A fortiori*, for all $m > 1$, Y_n is not true. So, Y_1 is true. But, by (*), Y_1 is not true. This contradiction shows that Y_0 is not true. So, by (A), for some $m > 0$, Y_m is true. Let k be a witness. So, Y_k is true. Again, from (A), for all $m > k$, Y_m is not true (**). *A fortiori*, for all $m > k + 1$, Y_m is not true. So, by (A), Y_{k+1} is true. But, from (**), Y_{k+1} is not true. Contradiction.

The Uniform Homogeneous Yablo Principle has the logical form:

For all x, $f(x)$ is P if and only if, for all $y > x$, $f(y)$ is not P.

which may be further schematized,

(B) Uniform Homogeneous Yablo Scheme

$$\forall x[\varphi(x) \leftrightarrow \forall y(\psi(y, x) \rightarrow \neg\varphi(y))].$$

Analysis of the inconsistency derivation shows that the Uniform Homogeneous Yablo Scheme is inconsistent with the following axioms for the relation symbol ψ:

(a) $\forall x \exists y \psi(y, x)$,

(b) $\forall x \forall y \forall z(\psi(x, y) \wedge \psi(y, z) \rightarrow \psi(x, z))$.

First, note that the above inconsistency argument does *not* establish that the *list* of Yablo sentences is paradoxical. It establishes only that the stronger *Uniform* Homogeneous Yablo Principle is inconsistent, and this is analogous to the trivial demonstration that the statement 'λ is true if and only if λ is not true' is inconsistent. Second, the inconsistency of the Uniform Homogeneous Yablo Principle has nothing to do with *truth*, for its inconsistency arises irrespective of what φ means: other than the Yablo scheme itself (B) and the auxiliary axioms (a), (b), no specific axioms for φ are used in the deduction of the inconsistency.

Third, and importantly, although the *Uniform* Homogeneous Yablo Scheme is inconsistent, we can demonstrate that its associated *set of numerical instances* is consistent. To see this, let L_F be the extension of the first-order language L of arithmetic, augmented with a primitive monadic predicate symbol F. Consider the theory $PA \cup \{\forall x[F(x) \leftrightarrow \forall y > x \neg F(y))]\}$. This theory is inconsistent, as shown above.

But consider instead the theory, $PA_F = PA \cup \{F(\underline{n}) \leftrightarrow \forall y > \underline{n} \neg F(y) : n \in \omega\}$. Then, PA_F is ω-inconsistent. For, consider $PA_F \cup \{F(\underline{n})\}$, for any $n \in \omega$. This implies $\forall y > \underline{n} \neg F(y)$, and thus $\forall y > \underline{n+1} \neg F(y)$, and thus both $F(\underline{n+1})$ and $\neg F(\underline{n+1})$. This contradiction shows that,

(a) $\quad PA_F \vdash \neg F(\underline{n})$, for all $n \in \omega$.

And thus, $PA_F \vdash \exists y > \underline{n}\, F(y)$, for all $n \in \omega$. So,

(b) $\quad PA_F \vdash \exists y F(y)$.

So, PA_F is ω-inconsistent.

This implies that no expansion of the *standard* model $\mathbf{N}$ of arithmetic satisfies PA_F. However, PA_F has a non-standard model. For let $\mathbf{M} \models PA$ be a non-standard model, and pick a non-standard element b.[4] Since b is non-standard, b is $>^{\mathbf{M}}$ than any 'standard' element of $\mathbf{M}$. Let $X = \{b\}$, where X is an interpretation of the predicate F and let $(\mathbf{M}, X)$ be the expanded model for the language L_F. Since b is non-standard, we have $(\mathbf{M}, X) \models \neg F(\underline{n})$, for all $n \in \omega$. Also, since X is non-empty, we have $(M, X) \models \exists y F(y)$. Since the witness is b, which is $>^{\mathbf{M}}$ than any 'standard' number in $\mathbf{M}$, we have $(\mathbf{M}, X) \models \exists y > \underline{n} F(y)$, for all $n \in \omega$. Thus, for all $n \in \omega$, $(\mathbf{M}, X) \models F(\underline{n}) \leftrightarrow \forall y > \underline{n} \neg F(y)$. So, $(\mathbf{M}, X) \models PA_F$.

This shows that the set of numerical instances of the Uniform Homogeneous Yablo Principle is consistent. It follows that the derivation of inconsistency in relation to the Yablo's Paradox must use the *Uniform* Homogeneous Yablo Principle, and not the set of its instances.[5] In particular, the list of Yablo sentences is not strictly paradoxical. Rather, the list might be called 'ω-paradoxical', in the sense that it is unsatisfiable on the standard model $\mathbf{N}$ of arithmetic.

2. FORMALIZING YABLO'S PARADOX

Graham Priest (1997) gave a natural formalization of the Yablo paradox. Priest showed how to construct the 'Yablo formula' $Y(x)$, such that,

($Y(x)$) For any number $y > x$, y does not satisfy $Y(x)$.

This equivalence is uniform. One cannot express this easily without abusing use/mention somewhat, but what it involves is roughly,

$$\forall x(Y(x) \leftrightarrow \text{for any number } y > x,\, y \text{ does not satisfy '}Y(x)\text{'})$$

Priest's main point is that it looks as if there is a subtle form of self-reference involved in Yablo's Paradox after all. For the Yablo formula $Y(x)$ now explicitly refers to *itself*. Technically, the 1-place Yablo predicate $Y(x)$ is a uniform fixed-point of the 2-place predicate 'No number larger than x satisfies z'.

We can formalize the paradox in the language L_T of arithmetic augmented with a primitive truth predicate T. The uniform diagonalization theorem implies the existence of a Yablo formula $Y(x)$ (containing T, and with a free variable x) such that,

(1) $PA \vdash \forall x(Y(x) \leftrightarrow \forall y > x \neg T(\ulcorner Y(\dot{y}) \urcorner)$.[6]

Each individual Yablo sentence Y_n is then $Y(\underline{n})$, where $\underline{n}$ is a numeral. Let us isolate this statement:

(C) The Uniform Fixed-Point Yablo Principle

$$\forall x(Y(x) \leftrightarrow \forall y > x \neg T(\ulcorner Y(\dot{y}) \urcorner)$$

The sub-formula $T(\ulcorner Y(\dot{y}) \urcorner)$ means 'the result of substituting the numeral of the number y for all free variables in the formula $Y(x)$ is true'. And this means 'y *satisfies* the formula $Y(x)$'. We can symbolize this as $S(\ulcorner Y(x) \urcorner, y)$. So, equivalently, we have:

(2) $PA \vdash \forall x(Y(x) \leftrightarrow \forall y > x \neg S(\ulcorner Y(x) \urcorner, y)$.

So, the Yablo formula $Y(x)$ is a fixed point (w.r.t. the free variable z) of the 2-place formula $\forall y > x \neg S(z, y)$, meaning 'no number larger than x satisfies z'.

We know that the Uniform Homogeneous Yablo Principle is inconsistent. To get from the Uniform Fixed-Point Yablo Principle to the homogeneous one, we need some kind of truth-theoretical principle. And naïvely, one might expect the Uniform Fixed-Point Yablo Principle to be inconsistent with the associated *local* disquotation principles,

(D) The Local Arithmetic Disquotation Scheme

$$T(\ulcorner \varphi \urcorner) \leftrightarrow \varphi, \text{ with } \varphi \in \text{Sent}(L).$$

This implements the equivalence of $T(\ulcorner \varphi \urcorner)$ and φ, when φ is an *arithmetic* sentence, lacking the truth predicate T.[7]

(E) The Local Yablo Disquotation Principle

$$T(\ulcorner Y_n \urcorner) \leftrightarrow Y_n, \text{ with } n \in \omega.$$

Let PA_Y be the theory $PA \cup (D) \cup (E)$. Then, PA_Y is ω-inconsistent. For $PA \vdash Y(\underline{n}) \leftrightarrow \forall y > \underline{n} \neg T(\ulcorner Y(\dot{y}) \urcorner)$, for all $n \in \omega$. So, $PA_Y \vdash T(\ulcorner Y(\underline{n}) \urcorner) \leftrightarrow \forall y > \underline{n} \neg T(\ulcorner Y(\dot{y}) \urcorner)$, for all $n \in \omega$. Let $F(x)$ be the formula $T(\ulcorner Y(\dot{x}) \urcorner)$. So, $PA_Y \vdash F(\underline{n}) \leftrightarrow \forall y > \underline{n} \neg F(y)$, for all $n \in \omega$. This is ω-inconsistent, by our earlier result that PA_F is ω-inconsistent.

Again, this implies that no expansion of the *standard* model **N** of arithmetic satisfies PA_Y. However, despite its ω-inconsistency, we can show that PA_Y is consistent. Indeed, PA_Y has a non-standard model. To prove this, select any non-standard model $\mathbf{M} \models PA$. Let # be a Gödel coding of L_T-sentences into the initial segment of 'standard' numbers in **M**. We need to define an expansion $(\mathbf{M}, E)$, where E is the interpretation of the truth predicate T, such that $(\mathbf{M}, E) \models PA_Y$. Let E_0 be $\{\#\varphi : \mathbf{M} \models \varphi$ and $\varphi \in \text{Sent}(L)\}$. Then E_0 contains only codes of arithmetic sentences. Let $t(y)$ be the term $\ulcorner Y(\dot{y}) \urcorner$. Let $t^{\mathbf{M}}$ be the function it denotes in **M**. Because **M** is non-standard, there exist non-standard elements c, b such that $c = t^{\mathbf{M}}(b)$. (From the non-standard model's viewpoint, c is the bth Yablo sentence.) Now, let $E = E_0 \cup \{c\}$. In particular, we have $(\mathbf{M}, E) \models_\sigma y > \underline{n} \wedge T(\ulcorner Y(\dot{y}) \urcorner)$, for all $n \in \omega$ (where σ is an **M**-assignment such that $\sigma(y) = c$). So $(\mathbf{M}, E) \models \exists y > \underline{n} T(\ulcorner Y(\dot{y}) \urcorner)$, for all $n \in \omega$. It is easy to see that $(\mathbf{M}, E)$ satisfies the Local Arithmetic Disquotation Scheme, since each element of E is either the code of an arithmetic truth in **M**, or is c. So, we need to prove that $(\mathbf{M}, E)$ satisfies the local Yablo disquotation principle. That is: $(\mathbf{M}, E) \models Y_n \leftrightarrow T(\ulcorner Y_n \urcorner)$, for each $n \in \omega$. The code of each Yablo sentence Y_n is neither the code of an arithmetic sentence, nor is identical to c. So, for any Y_n, we have $\#Y_n \notin E$. So, $(\mathbf{M}, E) \models \neg T(\ulcorner Y_n \urcorner)$, for each $n \in \omega$. Next, recall that $PA \vdash Y_n \leftrightarrow \forall y > \underline{n} \neg T(\ulcorner Y(\dot{y}) \urcorner)$, for each $n \in \omega$. So, $(\mathbf{M}, E) \models Y_n \leftrightarrow \forall y > \underline{n} \neg T(\ulcorner Y(\dot{y}) \urcorner)$, for each $n \in \omega$. But, $(\mathbf{M}, E) \models \exists y > \underline{n} T(\ulcorner Y(\dot{y}) \ulcorner)$. So, $(\mathbf{M}, E) \models \neg Y_n$, for each $n \in \omega$. So, $(\mathbf{M}, E) \models Y_n \leftrightarrow T(\ulcorner Y_n \urcorner)$, for each $n \in \omega$.[8]

Since PA_Y is consistent, it follows that the (inconsistent) Uniform Homogeneous Yablo Principle is not provable in PA_Y. Each *numerical instance* $T(\ulcorner Y_n \urcorner) \leftrightarrow \forall x > \underline{n} \neg T(\ulcorner Y(\dot{x}) \urcorner)$, with $n \in \omega$, is provable. But the generalization is *not* provable. It follows that, in order to deduce an inconsistency from the Uniform Fixed-Point Yablo principle, one needs (an instance of) the *uniform disquotational T-scheme*.

More exactly, one needs the truth-theoretic axiom,

(F) The Uniform Yablo Disquotation Principle

$$\forall x(T(\ulcorner Y(\dot{x}) \urcorner) \leftrightarrow Y(x)).$$

Note that this may be equivalently stated (by definitions) using satisfaction,

$$\forall y(S(\ulcorner Y(x) \urcorner, y) \leftrightarrow Y(y)).$$

This uniform truth-theoretic principle is stronger than the Local Yablo Disquotation Principle (E). In fact, as expected, the theory $PA \cup \{\forall x(Y(x) \leftrightarrow T(\ulcorner Y(\dot{x}) \urcorner))\}$ is inconsistent. For PA proves $\forall x(Y(x) \leftrightarrow \forall y > x \neg T(\ulcorner Y(\dot{y}) \urcorner)$. Using the uniform Yablo disquotation principle (F), we can semantically ascend, and prove $\forall x(T(\ulcorner Y(\dot{x}) \urcorner) \leftrightarrow \forall y > x \neg T(\ulcorner Y(\dot{y}) \urcorner)$. The logical form of this is $\forall x(\varphi(x) \leftrightarrow \forall y > x \neg \varphi(y))$, the Uniform Homogeneous Yablo Scheme, which is inconsistent.

3. SUMMARY

We have isolated the following six statements and/or schemes:

(A) The Uniform Homogeneous Yablo Principle

For all n, Y_n is true if and only if, for all $m > n$, Y_m is not true.

(B) Uniform Homogeneous Yablo Scheme

$$\forall x[\varphi(x) \leftrightarrow \forall y(\psi(y, x) \rightarrow \neg \varphi(y))].$$

Auxiliary axioms:

$$\forall x \exists y \psi(y, x),$$
$$\forall x \forall y \forall z(\psi(x, y) \wedge \psi(y, z) \rightarrow \psi(x, z)).$$

(C) The Uniform Fixed-Point Yablo Principle

$$\forall x(Y(x) \leftrightarrow \forall y > x \neg T(\ulcorner Y(\dot{y}) \urcorner)$$

(D) The Local Arithmetic Disquotation Scheme

$$T(\ulcorner \varphi \urcorner) \leftrightarrow \varphi, \text{ with } \varphi \in \text{Sent}(L).$$

(E) The Local Yablo Disquotation Principle

$$T(\ulcorner Y_n \urcorner) \leftrightarrow Y_n, \text{ with } n \in \omega.$$

(F) The Uniform Yablo Disquotation Principle

$$\forall x(T(\ulcorner Y(\dot{x}) \urcorner) \leftrightarrow Y(x)).$$

And their properties may be summarized as follows:

1. The scheme (B) is inconsistent (with the auxiliary axioms for the order).
2. The set of numerical instances of (B) is consistent, albeit ω-inconsistent.
3. (A) implies (B). So (A) is inconsistent.
4. (C) + (D) + (E) is consistent, albeit ω-inconsistent.
5. (C) + (F) implies (A). So (C) + (F) is inconsistent.

The central point is that, when examining the Yablo Paradox, if one demands a *formal* inconsistency, rather than an ω-inconsistency, one must use both the *uniform* fixed-point principle and the *uniform* disquotation principle. The local versions (sets of numerical instances) of these principles are consistent, and satisfiable on non-standard models.

NOTES

[1] And similarly for 'loopy' semantical paradoxes, such as:

(μ) ν is true.

(ν) μ is not true.

Formally, the set of biconditionals $\{\mu \leftrightarrow T(\ulcorner \nu \urcorner), \nu \leftrightarrow \neg T(\ulcorner \mu \urcorner)\}$ is inconsistent with the associated disquotational T-sentences $\{T(\ulcorner \mu \urcorner) \leftrightarrow \mu, T(\ulcorner \nu \urcorner) \leftrightarrow \nu\}$.

[2] One can 'formalize' Yablo's paradox *without mentioning truth at all*, as Thomas Forster notes (Forster 1996). Just use infinitary propositional logic with a denumerable list of sentence letters $p_1, p_2, \ldots$ Then take as axioms each formula:

$$A_n =_{\text{df}} p_n \leftrightarrow \wedge\{\neg p_k : k > n\}$$

where $\wedge \{\neg p_k : k > n\}$ is an infinitary formula, the infinite conjunction of formulae $\neg p_k$, with $k > n$. It is easy to show that the set $\Gamma = \{A_n : n \in \omega\}$ is unsatisfiable. Indeed, the assumption that Γ is satisfiable reduces (in the meta-theory) to the Uniform Homogeneous Yablo Principle discussed below. As Forster points out, this is a nice illustration of the failure of the compactness theorem for infinitary languages (although Γ is unsatisfiable, each finite subset of Γ is satisfiable).

[3] This principle is 'homogeneous' in the sense that it has the truth predicate on both sides of the biconditional.

[4] See Kaye 1991 for a detailed explanation of the properties of non-standard models of PA. The only property we need is the existence of non-standard 'infinite' elements, which are 'larger' than all the 'standard' numbers in the model.

[5] At least if we ignore infinitary logic. If we formalize Yablo's Paradox using infinitary logic, the Uniform Homogeneous Yablo Principle reappears in the meta-theory (see footnote 2 above).

[6] Here we use Feferman's 'dot notation', which allows us to 'quantify into' quotations. When $(\dot{x})$ appears inside a quotation term, it is defined such that the term $\ulcorner \varphi(\dot{x}) \urcorner$ is an *open* function term, with x free, meaning 'the result of substituting the numeral of the number x for all free variables in the formula φ'. More exactly, $\ulcorner \varphi(\dot{x}) \urcorner$ can be defined as $sub(num(x), \ulcorner \varphi(x) \urcorner)$, where the function term $sub(x, y)$ means 'the result of substituting x for all free variables in y' and $num(x)$ means 'the numeral of x'. In contrast, note that $\ulcorner \varphi(x) \urcorner$) is a *closed* quotation term, in which the variable x is *not* free.

[7] The *full* disquotation scheme is trivially inconsistent, since $\lambda \leftrightarrow \neg T(\ulcorner \lambda \urcorner)$ is provable in syntax. So, in order to study the situation here, we must concentrate on weaker truth-theoretic principles, so we can 'isolate' the reasoning to inconsistency. The local arithmetic disquotation scheme is very weak. In fact, it is conservative over almost any 'reasonable' base theory in L.

[8] The above model-theoretic proof in fact establishes that PA_Y is a *conservative* extension of PA. For suppose that PA does not prove an arithmetic sentence φ. Then, there is a model of $PA \cup \{\neg\varphi\}$. Indeed, there is an *non-standard* such model **M**. Then, by the above proof, **M** may be expanded to a model $(\mathbf{M}, E)$ of PA_Y. Since φ lacks the T-symbol, φ is still false in $(\mathbf{M}, E)$, and so PA_Y does not prove φ. By contraposition, if PA_Y proves φ, then PA proves φ too.

REFERENCES

Forster, T.: 1996, 'Yablo's Paradox', available at www.dpmms.cam.ac.uk/~tf/
Kaye, R.: 1991, *Models of Peano Arithmetic*, Oxford University Press, Oxford.
Priest, G.: 1997, 'Yablo's Paradox', *Analysis* **57**, 236–242.
Yablo, S.: 1993, 'Paradox without Self-Reference', *Analysis* **53**, 251–252.

University of Cambridge
Faculty of Philosophy
Sidgwick Avenue
CB3 9DA Cambridge
U.K.
E-mail: jjk32@cam.ac.uk