The structural collapse approach reconsidered

Ignacio Ojea

UBA - CONICET

Abstract

I will argue that Roy Cook's reformulation of Yablo's Paradox in the infinitary system D is a genuinely non-circular paradox, but for different reasons than the ones he sustained. In fact, the first part of the job will be to show that his argument regarding the absence of fixed points in the construction is insufficient to prove the non-circularity of it; at much it proves its non-self referentiality. The second is to reconsider the structural collapse approach Cook rejects, and argue that a correct understanding of it leads us to the claim that the infinitary paradox is actually non-circular.

1 Fixed points

As previously mentioned, Cook presents an infinitary system D for L_P including infinitary rules o f derivations, and shows that the (infinitary) Yablo's Paradox presented in Barrio's previous contribution leads to contradictions. Furthermore, he proves THEOREM 2.4.3, where he claims that the absence of weak fixed points is enough to show the non-circularity of the construction. I believe this is false.

Consider the set of sentence's names $\{S_1, S_2, S_3\}$ and the following denotation function δ:

- $\delta(S_1) = F(S_2)$.
- $\delta(S_2) = F(S_3)$
- $\delta(S_3) = F(S_1)$.

It is clear that the pair $< \{S_1, S_2, S_3\}, \delta >$ is closed and it is easy to check that we can derive a contradiction from those statements (in δD, the same for Yablo's paradox). Furthermore, it seems natural to claim that ***there is a circularity involved in*** $\{S_1, S_2, S_3\}$**.** After all, S_1 claims the falsity of S_2, S_2 claims the falsity of S_3 and S_3 claims the falsity of S_1. Nevertheless, ***neither*** S_1***,*** S_2 ***nor*** S_3 ***are weak fixed points of*** $< \{S_1, S_2, S_3\}, \delta >$, at least with Cook's definitions for D.

Let us start with the notion of fixed point under consideration:

- Φ is a weak sentential fixed point of $< \{S_\beta\}_{\beta \in B}, \delta >$ if and only if there is an $\alpha \in B$, a formula Ψ and statement name S_γ occurring in Ψ such that:
  - $\delta(S_\alpha) = \Phi$, and both
  - $\Phi \Rightarrow \Psi[S_\gamma/S_\alpha]$ and
  - $\Psi[S_\gamma/S_\alpha] \Rightarrow \Phi$ are theorems in D.

Since D is sound over the assignment semantics previously presented, to prove that there are no fixed points in $\{S_1, S_2, S_3\}$ I only need to give countermodels to the conditionals in the definition. This is in fact the very same strategy Cook uses to show THEOREM 2.4.3. I will only do it for S_1, the proofs for S_2 and S_3 are completely analogous.

Suppose $\delta(S_1)$ is a weak fixed point. Given that $\delta(S_1) = F(S_2)$, the formation rules for the extended L_P language used in the deductive system guarantee that $\Psi[S_\gamma, S_1]$ must take one of two forms, $T(S_1)$ or $(\Phi_1 \wedge \Phi_2 \wedge ...F(S_1)... \wedge \Phi_n \wedge \Phi_{n+1})$, where $T(x)$ works like a truth predicate. So we can only have:

a. $F(S_2) \Leftrightarrow T(S_1)$, or

b. $F(S_2) \Leftrightarrow (\Phi_1 \wedge \Phi_2 \wedge ...F(S_1)... \wedge \Phi_n \wedge \Phi_{n+1})$.

But now consider the following denotation function δ' and acceptable assignments σ_a and σ_b over it:

- $\delta'(S_1) = F(S_2) \wedge F(S_3)$, $\delta'(S_2) = F(S_1) \wedge F(S_3)$ and $\delta'(S_3) = F(S_2) \wedge F(S_1)$.
- $\sigma_a(S_1) = f$, $\sigma_a(S_2) = f$ and $\sigma_a(S_3) = t$.
- $\sigma_b(S_1) = t$, $\sigma_b(S_2) = f$ and $\sigma_b(S_3) = f$.

It is now easy to check that σ_a falsifies the a. biconditional, and σ_b falsifies b. biconditional. Hence neither biconditional is a theorem of D. The same can be proved for S_2 and S_3.

In conclusion, the absence of weak fixed points can be -at much- evidence for the non self-referentiality of the construction, but not for its non circularity.

2 The structural collapse approach reconsidered

There is, nevertheless, a way to argue in favor of the non-circularity of Cook's construction: by using the structural collapse approach. I will avoid presenting it because it was fully explained previously.

Cook offers two central arguments against it. The first one states, in a nutshell, that "the analogy -between statements and sets- can be, at most, suggestive (..) the identity of the characteristic sets does not imply the identity of the statements of which they are characteristic". Hence any information we might get by comparing characteristic sets may not be accurate regarding the statements of which they are characteristic. The second one runs like this: "even were the structural collapse account successful, it is not clear that it would provide what its defender presumably desires: An argument that the Yablo paradox involves circularity of the sort that can be blamed for the paradoxes. In other words, even if the structural collapse account entails the circularity of the Yablo paradox, it does not entail that the Yablo paradox suffers from a sort of circularity which can be blamed for the paradoxes".

I believe both arguments assume that the structural collapse approach fails in mapping sets of statements into (non-well-founded) sets in a way that *the referential structure of the statements is **isomorphic** to the membership relation of the sets.* To put it the other way around: If all the referential -hence semantic- properties of each (set of) statement(s) have a counterpart in membership properties of the characteristic set(s) and vice versa, and the characteristic set(s) of some (set of) statement(s) is clearly circular -under some understanding of this notion-, then it *must* be the case that the (set of) statement(s) is circular; this is exactly what the isomorphism guarantees. So the "stronger" the isomorphism is -in the sense that preserves all the referential and semantic properties over the mapping- the stronger the analogy is. In particular, if we can ensure that properties such as paradoxicality have some kind of counterpart in the non-well-founded sets, and the characteristic set of Yablo's paradox is both circular and paradoxical in that sense, then "the Yablo paradox suffers from a sort of circularity which can be blamed for paradox".

But the reciprocal will also be the case: if the mapping genuinely preserves all semantic properties and the characteristic set of Yablo's paradox is both paradoxical and non-circular; then we have found a non-circular paradox. I will argue this is the case.

Let us start considering the mapping operation. It is a composition of two operations, a mapping from sentences to direct graphs and a mapping from graphs to (possibly non-well-founded) sets using some set theory.

The connection between graph theory and L_P is straightforward. Let $D_\delta(S_i)$ be the set of all sentence names that appear on the δ-denotation of S_i. So for example if A is a set of indexes and $\delta(S_i) = \bigwedge\{F(S_\alpha) : \alpha \in A\}$, then $D_\delta(S_i) = \{S_\alpha : \alpha \in A\}$. A pair with set of sentence names and a denotation function, $< \{S_\beta\}_{\beta \in B}, \delta >$, is *closed* iff for any α, β, if $\beta \in B$ and $S_\alpha \in D_\delta(S_\beta)$, then $\alpha \in B$. All pairs considered here will be closed in this sense.

An assignment is a function $\sigma : C \to \{t, f\}$ and it is *acceptable* over a denotation function δ if and only if for every $\beta \in B$: $\sigma(S_\beta) = t$ iff, for all $S_\alpha \in D_\delta(S_\beta), \sigma(S_\alpha) = f$. Given a set of indexes B and $< \{S_\beta\}_{\beta \in B}, \delta >$ closed, let $Dep_\delta(\{S_\beta\}_{\beta \in B}) = \{< S_\alpha, S_\gamma >: \alpha, \gamma \in B; S_\gamma \in D_\delta(S_\alpha)\}$ be the overall dependence relation over all the sentences' names considered. Then the pair:

- $< \{S_\beta\}_{\beta \in B}, Dep_\delta(\{S_\beta\}_{\beta \in B}) >$

is a serial directed graph. It is also easy to check that given some sentence name S_α in the language, the subgraph of $< \{S_\beta\}_{\beta \in B}, Dep_\delta(\{S_\beta\}_{\beta \in B}) >$ that starts with S_α is an accessible pointed directed serial graph.

Furthermore, we can represent assignments of truth values to sentences in L_P relative to a denotation function δ with *coloring* of graphs. A coloring of a graph is nothing more than an assignment of a color to each node in the graph. We shall consider only two colors, turquoise and fuchsia, and we shall impose the following conditions on acceptable colorings:

- Given a serial, directed graph $< N, E >$, a coloring of $< N, E >$ is *acceptable* if and only if, for any $n \in N$: n is colored turquoise if and only if, for any node m such that $< n, m > \in E$, m is colored fuchsia.

Any assignment σ that is acceptable on $< \{S_\beta\}_{\beta \in B}, \delta >$ induces a coloring on the graph $< \{S_\beta\}_{\beta \in B}, Dep_\delta(\{S_\beta\}_{\beta \in B}) >$ (and vice versa) in which (a) S_β is colored turquoise if and only if $\sigma(S_\beta) = t$, and (b) S_β is colored fuchsia if and only if $\sigma(S_\beta) = f$.

So *there is a one to one correspondence between sets of sentences in L_P and graphs* ***that preserves semantic properties.*** If a set of sentences is paradoxical, it does not admit an acceptable assignment; so its associate graph does not admit an acceptable coloring. In this sense, we are able to characterize a graph as paradoxical whenever it does not admit an acceptable coloring.

Cook observes there are mainly four ways of mapping graphs into non-well founded sets. In order to take a closer look to them, some definitions are required. Say U is the universe of non-well-founded sets and $< N, E, p >$ is an accessible pointed directed graph -APG- (where N is the set of nodes, $E \subseteq N^2$ and p is the origin of the graph), then:

- a function $f : N \to U$ is a *decoration* of the APG $< N, E, p >$ iff for any $n_1, n_2 \in N$: $f(n_1) \in f(n_2)$ iff $< n_1, n_2 > \in E$.
- A decoration f on an APG $< N, E, p >$ is a *picture* of $S \in U$ iff $f(p) = S$.
- A decoration f on an APG is *exact* iff, for any distinct $n_1, n_2 \in N$, $f(n_1) \neq f(n_2)$.
- An APG $< N, E, p >$ is an exact picture of S iff there is a decoration f such that (a) f is a picture of S, and (b) f is exact.

Each set can have at most one exact picture (up to isomorphism), and each graph can be the exact picture of at most one set. Further, all four candidates for anti-foundation axiom agree that any APG is a picture of a set. Where there they disagree is in terms of which graphs provide exact pictures of sets. The strongest candidate for an anti-foundation axioms (strongest in the sense that it provides the weakest criterion for a graph providing an exact picture, and thus allows for the greatest variety of distinct non-well-founded sets) is the Boffa Anti-foundation Axiom (or BAFA) BAFA states, in essence, that every extensional APG has is an exact picture, where extensional is defined as follows:

- An APG $< N, E, p >$ is *extensional* iff, for any $n_1, n_2 \in N$: If for all $m \in N$, $< n_1, m > \in E$ iff $< n_2, m > \in E$, then $n_1 = n_2$.

In this case, the 'Liar' set Ω and 'Yablo's' under Cook's infinitary reformulation are distinct sets, since they correspond to distinct extensional graphs. Thus, if BAFA is the correct, or best, account of non-well-founded set theory, then the structural collapse account of circularity does not entail the circularity of the Yablo paradox.

The second anti-foundation axiom is FAFA (whose addition to ZFC-foundation results in Finzler-Aczel set theory). FAFA states that an APG $< N, E, p >$ must not only be extensional, but must also be isomorphism-extensional, if it is to be an exact picture of a set:

- Given an accessible pointed graph $< N, E, p >$, the sub-APG induced by m (where $m \in N$) is $< N^*, E^*, m >$ where (a) $N^* = \{q : \text{there is a path from } m \text{ to } q\} \cup \{m\}$, and (b) $E^* = E \cap (N^* x N^*)$.
- An accessible pointed directed graph $< N, E, p >$ is *isomorphism-extensional* iff, for any $n_1, n_2 \in N$, if the sub-APG induced by n_1 is isomorphic to the sub-APG induced by n_2, then $n_1 = n_2$.

Intuitively, an APG is an exact picture of a set if no two sub-APGs are isomorphic. In particular, if we look at the APG induced by the infinitary version of Yablo's paradox, the subgraphs induced by each of the nodes are isomorphic. So that graph is not an exact picture of a set. On the other hand, the graph is a picture of the Liar set, since we can decorate each node with the Liar set Ω. As a result, in Finzler-Aczel set theory the characteristic set(s) corresponding to the Yablo paradox are not distinct from the characteristic set corresponding to the Liar paradox.

The two other theories, known as SAFA and AFA -for Scott and Aczel-, will not be treated here, since we have the following result. Letting G_Φ be the class of APGs that are exact pictures of sets relative to a particular (non-well-founded) set theory whose anti-foundation axiom is Φ:

- $G_{AFA} \subseteq G_{SAFA} \subseteq G_{FAFA} \subseteq G_{BAFA}$

Hence, the APG corresponding to Yablo's Paradox will not be an exact picture neither in AFA nor SAFA; and in both theories it will be a picture of the Liar set Ω.

In conclusion, it seems that in order to decide whether the paradox is circular we need to choose between non-well-founded set theories, since AFA, SAFA and FAFA regard it as structurally equivalent to the Liar and BAFA does not. But we do have a reason for choosing BAFA: it is the only theory that preserves semantic properties such as paradoxicality.

Consider the following sentences in L_P:

- $\delta(S_1) = F(S_2)$
- $\delta(S_2) = F(S_1)$

It is easy to define an acceptable assignment for $\{S_1, S_2\}$ over δ. We could use $\sigma_a(S_1) = t$ and $\sigma_a(S_2) = f$ or $\sigma_b(S_1) = f$ and $\sigma_b(S_2) = t$. Hence that set of sentences is not paradoxical, admits acceptable assignments. Analogously, we could construct two acceptable colorings for the associate graph of S_1 (and S_2), so the graph is also non-paradoxical.

Let us look now what happens with the different non-well-founded set theories. Note that the associate graph of S_1 is extensional: no two distinct nodes have the same children. Then in BAFA it is an exact picture of set distinct from the characteristic set of the Liar. This is what we need, *otherwise GAFA would identify a non-paradoxical sentence with a paradoxical one.*

On the contrary, in FAFA fails at this point. In the graph of S_1, the sub-APG induced by S_1 is isomorphic to the sub-APG induced by S_2, but $S_1 \neq S_2$. Hence the graph is not an exact picture of a set. But it can be decorated with the Liar's set Ω. As a result, in Finzler-Aczel set theory the characteristic set corresponding to S_1 is not distinct from the characteristic set corresponding to the Liar paradox. Using the previously mentioned fact that $G_{AFA} \subseteq G_{SAFA} \subseteq G_{FAFA} \subseteq G_{BAFA}$, we note that the graph of S_1 is not an exact picture in AFA and SAFA. And for the same reason as before, the characteristic set of S_1 is also Ω in those theories. In conclusion, *FAFA, SAFA and AFA do not preserve the semantic properties of the graphs (sentences)*; since they identify as structurally isomorphic a paradoxical sentence and a non-paradoxical sentence.

It may be argued that we do not *need* to preserve paradoxicality in order to guarantee that the mapping preserves circularity. That is true, but I believe it would undermine the whole structural collapse approach. Circularity is not problematic *per se*, it is only problematic when it leads to inconsistency. What the structural collapse must guarantee is for as to have some way to identify *the sort of circularity that leads to paradox.* No set theory that dissipates the difference between sentences that admit acceptable assignments and sentences will be useful in doing so.

So if my point is conceded, we should take BAFA to be the adequate theory for the structural approach account. Then the characteristic set of Yablo's paradox is distinct from the characteristic set of the Liar. Now: have we shown that Cook's reconstruction of Yablo's paradox is not circular? Certainly not. At this point we can only claim that it lacks of *some* well known sort of circularity, the one present in the Liar. But we did not show that is non-circular *simpliciter.* That task would require a characterization of the notion of circularity for non-well-founded sets, and it is beyond the scope of the present work. Nevertheless, some rough idea can be shortly presented.

Let us say that an APG $< N, E, p >$ is *circular* whenever there is at least one node $n \in N$ such that there is a *finite* path from n to n. And let us say a non-well-founded set is circular whenever its exact picture is circular. Analogously, a (set of) sentence(s) is circular if its characteristic set(s) is circular. Of course, this intuitive idea of circularity may be problematic in hard cases, but I take it to be simple and clear enough for the present purposes. Under the present definition, Yablo's paradox is not circular at all: in the graph associated with that set of sentences there is no node connected to itself by a finite path.