Explanation and Plenitude in Non-Well-Founded Set Theories[1]

Ross P Cameron
University of Virginia

This article has been accepted for publication in *Philosophia Mathematica* Published by Oxford University Press.

Abstract

Non-well-founded set theories allow set-theoretic exotica that standard ZFC will not allow, such as a set that has itself as its sole member. We can distinguish plenitudinous non-well-founded set theories, such as Boffa set theory, that allow infinitely many such sets, from restrictive theories, such as Finsler-Aczel or AFA, that allow exactly one. Plenitudinous non-well-founded set theories face a puzzle: nothing seems to explain the identity or distinctness of various of the sets they countenance. In this paper I aim to sharpen this puzzle, make clear who it does and does not apply to and, ultimately, to argue in favor of a plenitudinous theory like Boffa.

0 Introduction

This paper is an investigation of non-well-founded set theories, and how best to solve some puzzles that arise concerning non-well-founded sets and explanatory gaps and arbitrariness. The overall argument will be that while more restrictive non-well-founded set theories (like Finsler-Aczel or AFA set theory) might initially appear to do better in solving the puzzles, in fact a plenitudinous set theory like Boffa set theory does better overall. On our way we look at some interesting issues concerning explanation and arbitrariness, the individuation of sets, and semantic phenomena such as the truth-teller sentence and its like.

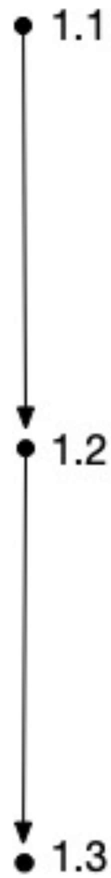
In section 1 I introduce three non-well-founded set theories: Boffa, Finsler-Aczel and AFA. In section 2, I discuss an argument I made in previous work that all three theories inevitably result in explanatory gaps. In section 3, I discuss two different ways of individuating sets and argue that, armed with these extra resources, the defender of Finsler-Aczel or AFA can avoid the problem I raised, but the defender of Boffa cannot. However, in section 4 I go on to argue that the problem is not fatal for Boffa either, drawing out some morals concerning plenitude and arbitrariness and making an analogy with some cases adjacent to the semantic paradoxes.

[1] Thanks to Elizabeth Barnes, Jc Beall, Roy Cook, Scott Dixon, Maegan Fairchild, Chris Menzel, Trenton Merricks, Daniel Nolan, Graham Priest, Adam Rieger, Jason Turner, Robbie Williams, an audience at Virginia Tech, and two anonymous referees and the editor of *Philosophia Mathematica* for helpful discussion.

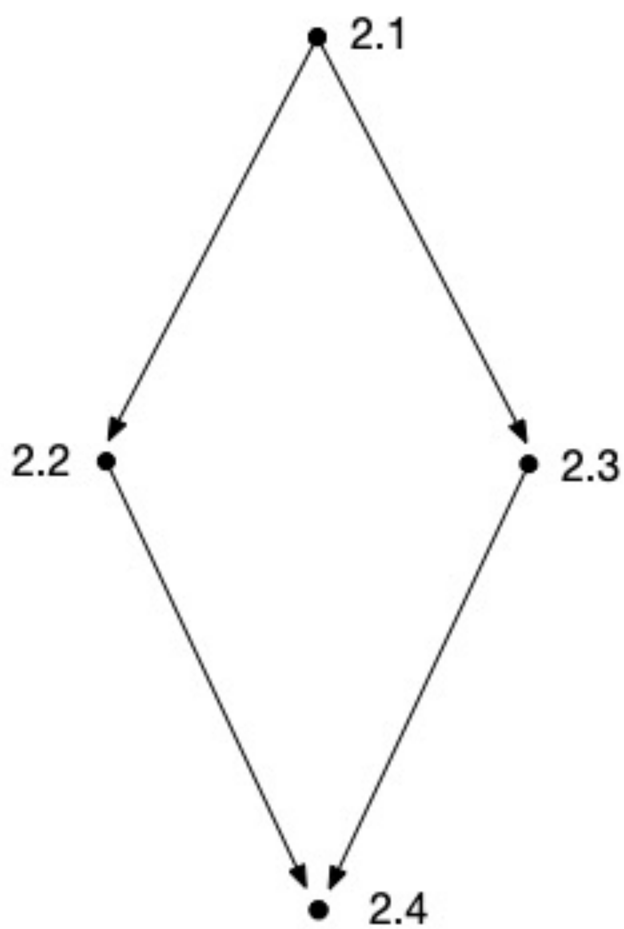
Finally, in section 5, I argue for Boffa set theory over its non-well-founded rivals, by arguing that their response to the problem I raised is the better one, especially once we also take into account impure sets.

1 Three Non-well-founded Set Theories

Let us start by putting four set theories on the map. As is familiar, I will use accessible pointed graphs (henceforth, just 'graphs') to represent sets. Each node on a graph represents a set; the topmost node (the point) represents the set that the graph as a whole represents, and an arrow going from node n to node m represents that the set represented by node n has as a member the set represented by node m. So for example the graph in Fig.1 represents the singleton of the singleton of the empty set, because node 1.3 represents the empty set (as it has no arrows coming out from it, so represents a set that has no members), node 1.2 represents the singleton of that set, and node 1.1 represents *its* singleton.



Some graphs represent a set in a special way: such that each node on the graph represents a distinct set (i.e. a set not represented by any other node on that graph). When a graph, g, represents a set, S, in this manner we say that g is an *exact* representation of S. The graph in Fig.1 is an exact representation of $\{\{\varnothing\}\}$, for no two of the three nodes represent the same set. Other graphs cannot be an exact representation of any set. For example, the graph in Fig.2 cannot be an exact representation of any set, given extensionality, because nodes 2.2 and 2.3 cannot represent distinct sets, and that is because those nodes have the same children: i.e. the nodes that are pointed to by an arrow coming from 2.2 and the nodes that are pointed to by an arrow coming from 2.3 are exactly the same. That means that the sets represented by nodes 2.2 and nodes 2.3 have the same members, in this case the empty set, which is represented by node 2.4. Hence, given extensionality, the sets represented by nodes 2.2 and nodes 2.3 are one and the same set, and so the graph in Fig.2 is not an *exact* representation of any set. The graph in Fig.2 can be an *inexact* representation of a set, one in which distinct nodes represent the same set. Indeed, it is an inexact representation of the very same set that is exactly represented by Fig.1: the singleton of the singleton of the empty set. But no graph that has two nodes that have exactly the same children can be an exact representation of any set, given extensionality (which I will take for granted throughout). For all that has been said so far, a single graph can be an exact representation of one set and an inexact representation of another, and a single graph can be an exact representation of more than one set. Whether this ever actually happens depends on the details of the particular set theory.



ZFC (or ZF: the axiom of choice will play no role in what follows) is a foundationalist set theory. Every pure set, according to ZFC, is ultimately 'built up' from the empty set. Hence, the iterative hierarchy of pure sets: at the bottom is the empty set, and at each higher level are the sets that can be constructed by taking as members sets already obtained at lower levels. In terms of the graphs, this means that every pure set that exists in the set-theoretic universe of ZFC can only be exactly represented by a graph that has a unique bottom node that has no arrows coming out of it: that is, any graph g that is an exact representation of a pure set in the ZFC universe has exactly one node, 0, such that no arrows come out from that node, and for every other node, n, on g, every path from n (that is, every chain of arrows you can follow coming out from n) terminates in node 0.

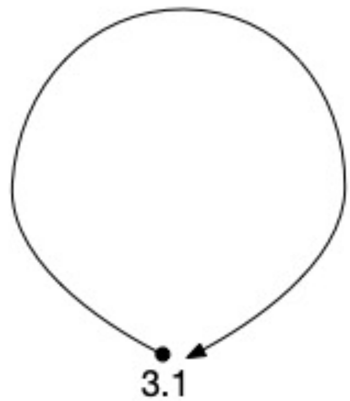
Non-well-founded set theories reject the foundationalist requirement of ZFC: that is, they reject that there is a foundational set - or even a foundational level of sets - from which all other sets are built. In terms of the graphs, this means that some sets belonging to the set-theoretic universe of a non-well-founded set theory can be exactly represented by a graph that has paths that do not terminate. That is, some graph, g, exactly represents some non-well-founded set, S, and there is at least one node, n, on g such that there is a chain from n that does not terminate, meaning that you can follow a chain of arrows originating from n, and that chain either goes on infinitely without end, or it doubles back on itself, creating a circular path.

We will look at three non-well-founded set theories: Boffa, Finsler-Aczel, and AFA.[2] They are differentiated by the rules they give on what a graph has to be like in order to be an exact representation of a set. On each theory, once you have settled those rules, you have determined the set-theoretic universe: a (pure[3]) set exists, according to the relevant theory, if and only if it is exactly represented by some graph that meets the relevant rules.

Boffa is the most liberal of the non-well-founded set theories we will look at: it says that a graph is an exact representation of a set iff no two nodes on that graph have the same children. That means that the graphs in Fig.3 and Fig.4 exactly represent some set whose existence is countenanced by Boffa set theory.

[2] For technical details see Aczel (1988) and Boffa (1969). For philosophical discussion of non-well-founded set theories see Cameron (2022, Ch.2), Incurvati (2014, 2020) and Rieger (2000).

[3] Unless I explicitly say otherwise, when I talk about sets, or the set-theoretic universe, I am talking about pure sets, and the universe of pure sets. We will have cause to consider impure sets in the final section.


3.1

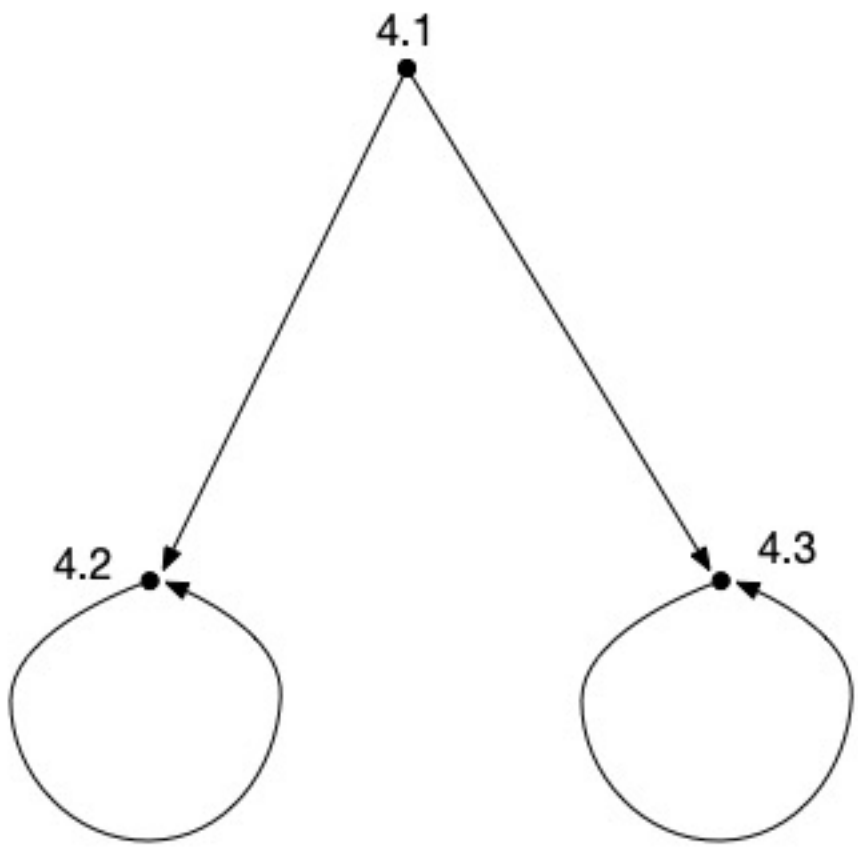
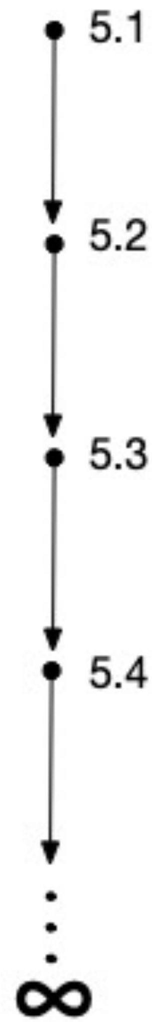

4.1
4.2
4.3

Fig.3 exactly represents a set that has exactly one member: itself. It is a self-singleton. Fig.4 exactly represents a set that has exactly two members, both of which are self-singletons. Fig.4 also *inexactly* represents a self-singleton, since we can consistently take each node to represent the very same set, one that has itself as its sole member; but this need not concern us, we are interested in what sets are exactly represented by these graphs, since that lets us understand the scope of the set-theoretic universe according to Boffa. And what is interesting about Fig.4 is not so much the set that the graph as a whole exactly represents, but rather that nodes 4.2 and 4.3 on that graph each represent distinct self-singletons. This means that on Boffa set theory, there can be two distinct sets, Ω1 and Ω2, each of which has itself as its sole member. That is: Ω1={Ω1} and Ω2={Ω2} and Ω1≠Ω2. And of course we could take the graph in Fig.4 and add another arrow coming out of node 4.1 to a new node, node 4.4, that has an arrow pointing just to itself; or another arrow from node 4.1 to node 4.5 that points just to itself . . . and so on. No matter how many arrows we have coming out of node 4.1 going to a node that points just to itself, the resulting graph will meet Boffa set theory's requirements for being an exact depiction of a set, since there will never be two nodes with the same children. Therefore, according to Boffa set theory, there are very many sets that have themselves as their sole member. In fact, according to Boffa set theory, there are proper class many self-singletons.



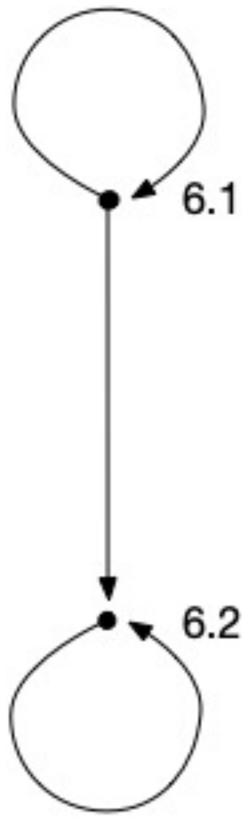
The graph in Fig.5 is also an exact depiction of a set according to Boffa set theory. No two nodes in the graph in Fig.5 share the same children, obviously, since each node 5.n has exactly one child, node 5.n+1. And so this graph exactly depicts a set that exists according to Boffa set theory: a singleton set whose sole member is a distinct singleton set whose sole member is yet another distinct singleton set whose sole member is yet another distinct singleton set . . . and so on, ad infinitum. Here we have an infinite chain of sets, each distinct from any of the others, and each of which has the next as its sole member.

Now let us consider Finsler-Aczel set theory. Finsler-Aczel set theory says that a graph is an exact depiction of a set iff it does not have distinct nodes with isomorphic sub-graphs beneath those nodes. This of course entails Boffa's demand that no graph that is an exact depiction has two distinct nodes that share the same children, since if two nodes on a graph point to exactly the same nodes, those two nodes will thereby determine identical, *a fortiori* isomorphic, sub-graphs. But Finsler-Aczel's demand is stronger than Boffa's and will rule out as exact depictions graphs that Boffa allows. For example, Finsler-Aczel will rule out Fig.4 as an exact depiction of a set, for nodes 4.2 and 4.3 of that graph have isomorphic sub-graphs. Likewise for any of the variants of Fig.4 that add yet another arrow coming out of node 1. Fig.3 is okay, though, even by the lights of Finsler-Aczel. And so while Finzler-Aczel set theory agrees with Boffa set theory that Fig.3 is an exact depiction of a set, and hence these set theories agree that there can be self-singleton sets, i.e. sets whose sole member is themselves, Finsler-Aczel rejects Boffa's *abundance* of self-singletons, and insists that there is exactly one such set, the set Ω such that Ω={Ω}. Fig.3, according to Finsler-Aczel, is an exact depiction of this unique self-singleton Ω, and Fig.4 (and its variants) is an inexact depiction of Ω, in which that unique self-singleton is represented by every node.

Nor is the graph in Fig.5 an exact depiction of a set, according to Finsler-Aczel set theory. On this graph, *any* two nodes have isomorphic sub-graphs, since the sub-graph under any given node is simply an image of the graph as a whole. So Finsler-Aczel set theory rejects the set with the infinitely descending membership chain that Boffa set theory allows. The graph in Fig.5 can be an *inexact* depiction of a set, according to Finsler-Aczel: indeed, it is once again an inexact depiction of the unique (according to Finsler-Aczel) self-singleton, Ω. Node 5.1 represents Ω, which has one member, Ω, which has one member, Ω, which has one member, Ω . . . and so on. Finsler-Aczel sees the graph in Fig.5 as nothing other than the infinite unpacking of the circular membership chain depicted by the graph in Fig.3; whereas Boffa set theory allows that in addition to inexactly depicting each of the proper class many self-singletons, the graph in Fig.5 exactly depicts an infinite chain of distinct singleton sets, each having as its member the next.

An even more restrictive set theory is AFA, which says that a graph is an exact depiction of a set just in case it is not possible to take two of its nodes to be depictions of the same set. This is a stronger restriction than that of both Boffa and Finsler-Aczel. It of course entails Boffa's very weak demand than in order to be an exact depiction of a set a graph must not have two nodes with the same children, since if two nodes have the same children it is possible - indeed, mandated, given extensionality - to take those nodes to depict the same set. And it entails

Finsler-Aczel's restriction, for if two nodes on a graph have isomorphic sub-graphs beneath them then it is always possible to take those two nodes to depict the same set by taking every node on the sub-graph under one of them to depict the same set as the corresponding node on the sub-graph under the other. The graph in Fig.6 gives a nice illustration of where Finsler-Aczel and AFA give different results.



According to Finsler-Aczel set theory, this graph could be an exact depiction of a set, S = {S,Ω}. That is, a set which has two members: itself, and the self-singleton Ω. It can also be an inexact depiction of Ω. Obviously we have no choice but to take node 6.2 to depict Ω, but when we then look at node 6.1 and see that it points to itself and to node 6.2, we can either interpret this as the set depicted by node 6.1 having two members, itself and Ω, or we can interpret it as saying twice that the set depicted by node 6.1 has Ω as a member, in which case the set depicted by node 6.1 just is Ω, the (unique, according to Finsler-Aczel) set which has itself as a member. But precisely because there is this choice - because it is *possible* to take nodes 6.1 and 6.2 to depict the very same set - AFA demands that this be the sole interpretation of the graph in Fig.6. That is, because it is possible to take this graph to be an inexact depiction of Ω, AFA rules out taking it to be an exact depiction of a doubleton set S = {S,Ω}. If a graph can be an inexact depiction of a set, says AFA, it cannot be an exact depiction of a set. Thus the alleged doubleton set S = {S,Ω} does not exist according to AFA, whereas it does exist according to Finsler-Aczel (and, *a fortiori*, Boffa).

2 Explanatory Gaps

We have looked at three non-well-founded set theories: Boffa, Finsler-Aczel, and AFA. In this section I start by looking at a puzzle I discussed in previous work[4] that faces Boffa set theory but which, on the face of it at least, Finsler-Aczel and AFA avoid. In that work, I argued that the puzzle in fact poses just as much of a problem for Finsler-Aczel and AFA as it does Boffa, and hence that there is no argument against Boffa set theory specifically. I now think I was wrong about this, and in the next section of this paper I will argue that in fact Finsler-Aczel and AFA have resources to respond to the puzzle that Boffa lacks. However, I do not come to bury Boffa set theory, and I will argue that a different response to the puzzle is available to the Boffa set theorist. In later sections I will argue that in fact the response from the Boffa set theorists is preferable, and hence that Boffa set theory is, at least in this respect, preferable to Finsler-Aczel and AFA.

The puzzle is that Boffa set theory seems to result in a certain kind of explanatory gap that yields unanswerable questions. Consider two of Boffa's self-singletons, Ω1 and Ω2. *Why* are they distinct? The only answer we can give, seemingly, is that they are distinct because they do not have the same members. But this 'explanation' is not really an explanation, for the distinctness of the sets in this case *just is* the distinctness of the members, since Ω1 and Ω2 are identical to their members. So we are attempting to explain the distinctness of the sets by appealing to the very fact that they are distinct. And that is no explanation at all: in no good explanation is the explanans the same as the explanandum. Similarly if we happen to pick out the same self-singleton twice. Suppose that 'Ω1' and 'Ω2' name the same self-singleton. Why is Ω1= Ω2? Again, the only answer that suggests itself is that they are identical because their members are identical. But again, this 'explanation' simply restates that which was to be explained, since Ω1 and Ω2 are their members. Our attempted explanation is no better than saying that Ω1= Ω2 because Ω1= Ω2. That is to say, it is no explanation at all.

So we cannot explain the identity or distinctness of pairs of self-singletons found in the Boffa set-theoretic universe by appealing to the identity or distinctness of their members. But there is nothing *else* that we could appeal to to explain these facts. Nothing is relevant to the identity or distinctness of these sets other than the identity or distinctness of their members. As I put it previously: "The beginning and end of the story, when it comes to the identity or distinctness of sets, is extensionality: this set is identical to/distinct from that set iff they have the same/different members."[5] I concluded in that earlier work that the defender of Boffa set theory must simply take facts concerning the identity or distinctness of self-singletons as explanatorily brute.

Once you see the recipe for the puzzle, you can see that such explanatory bruteness is abundant given Boffa set theory. Consider, e.g., Boffa's infinite sequence of sets, each of which is the

---

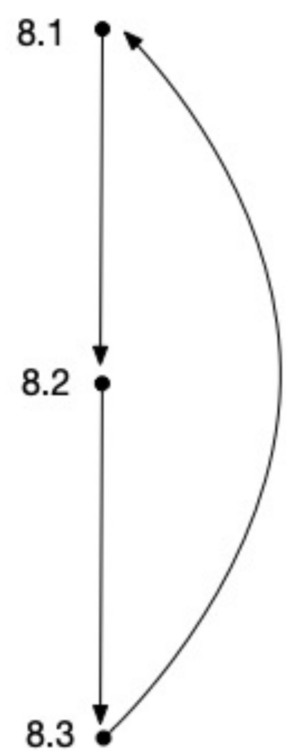
[4] Cameron (2022, Ch.2). The puzzle builds on some remarks made by Rieger (2000).

[5] Cameron (2022, p78).

singleton of the next, as depicted in Fig.5. There is not *one* such infinite sequence of sets, given Boffa set theory: just as with self-singletons, there are proper class many such sequences. Suppose, then, that we have a set S1, that has exactly one member, S3, that has exactly one member, S5, . . . and so on, *ad infinitum*. What makes S1 the very set S1? That it has the very set S3 as a member? But what makes that set the very set S3? That it has the very set S5 as a member? But what makes that set the very set S5? . . . This sequence of questions never stops. There is a sense that we never get a satisfying explanation, because the explanation simply raises a new demand for an explanation of the very same kind. Now, there is debate as to whether an infinite chain of explanatory demands and answers like this precludes explanation, and I don't aim to settle that here[6], but the explanatory gaps becomes especially vivid when we compare this sequence to another structurally identical one. Suppose we have a set S2, that has exactly one member, S4, that has exactly one member, S6, . . . and so on, *ad infinitum*. We can, of course, raise the same question about the individual sets in this sequence and end up in the same position of having an infinite sequence of questions and answers, and we can debate whether that precludes explanation. But where I think there is unquestionably an explanatory gap is when we ask the *contrastive* question: why is this set S1 the very set S1 as opposed to, say, S2? The only answer available to Boffa is that it has the very set S3 as a member, as opposed to S4. But why is that set that is a member of S1 the very set S3 as opposed to S4? Because it has the very set S5 as a member, as opposed to S6. But why . . . and so on. It seems that there can never be a good explanation as to why a set from one of these infinite sequences is that very set and not its counterpart set from the other infinite sequence (or its counterpart from yet another such sequence), because any such explanation simply presupposes that we're talking about the sets on that very infinite sequence as opposed to one of the others. Even if an infinite sequence of demands for explanation and answers to that demand can in principle be explanatory, this presupposition of what is to be explained rules out, I think, any good explanation of these contrastive facts.

[6] See Cameron (2022, esp. chapters 1&3) for discussion. Cf. Bliss (2013) and Priest (2014, p186).

Or consider a circular case of set membership, such as that depicted in Fig.7. Fig.7 depicts (assuming Boffa) two distinct sets, each of which has exactly one member, the other. It's like a set-theoretic version of Max Black's two spheres[7], with nothing to tell them apart other than their distinctness. The Boffa theorist has nothing illuminating to say about *why* they are distinct. In this case, of course, it's not that the sets are their members, but it's still the case that trying to explain the distinctness of the sets via the distinctness of their members is to try to explain a fact by itself, since each set in the pair has as its sole member the other set in that pair. So the distinctness of the sets still just is the distinctness of the members, and so their distinctness simply has to be taken as explanatorily brute by the Boffa theorist. Furthermore, in Boffa we can have two pairs of sets each of which is the singleton of the other - or three, or four, or as many as we like. So suppose we have four sets: S1={S2}, S2={S1}, S3={S4}, S4={S3}. Not only has Boffa no explanation for why each set is distinct from the other one in its pair, it has no explanation for the distinction between the pairs. Why, for example, are S1 and S2 the very pair of sets S1 and S2 and not the pair of sets S3 and S4? The only possible answer, seemingly, is that S1 is the very set S1 and not S3 because it has S2 as a member and not S4. But that explanation presupposes that it is the very set S2, and not S4, that belongs to S1. So why is that set the very set S2? Because it has the very set S1 as a member, as opposed to, e.g., the set S3 (which would make it S4 and not S2). But this set being the very set S1 is what we set out to explain in the first place. We've just gone in a circle, and that is no explanation at all.



[7] Black (1952).

Likewise with the set depicted in Fig.8. Why is the set depicted by node 8.1 that very set and not the set depicted by node 8.2, say? All we know about it is that it has one member, that has one member, that has as a member the very set we started with. But that is also true of the set depicted by node 8.2, and indeed by every node. So nothing explains why any of the four sets depicted on this graph is distinct from any other, and yet there must be four such distinct sets, for this graph exactly depicts some set, given Boffa set theory.

So if Boffa set theory is true, there seem to be many cases in which there are distinct sets where there is no explanation as to why those sets are distinct, because the only possible explanation as to the distinctness of the sets goes via the distinctness of their members, and just because of how those sets are structured this turns out to either be the very fact we are trying to explain, or to presuppose it, or to simply take us in an explanatory circle. There seems to be no option but for the Boffa set theorist to take some facts concerning the identity and distinctness of sets to be explanatorily brute.

It would be natural to think that this is a distinctive puzzle for Boffa. After all, the problem seemingly arises precisely because of the abundant set-theoretic universe that Boffa's liberal rules allow. Boffa set theory must take facts concerning the identity or distinctness of self-singletons as explanatorily brute *because* it allows multiple self-singletons. Finsler-Aczel and AFA, by contrast, only allow one self-singleton. According to those more restrictive set theories, if there if a self-singleton, Ω1, and a self-singleton, Ω2, then Ω1=Ω2, because there only is one self-singleton. The question as to why distinct self-singletons are distinct doesn't arise, because there *are no* distinct self-singletons. Likewise, there is no infinite sequence of distinct sets, each of which is the singleton of the other, let alone infinitely many such sequences of sets. Nor are there two distinct sets each of which is the singleton of the other. Finsler-Aczel and AFA seem to avoid the problem simply by restricting their set-theoretic universes in such a way that the sets that lead to the explanatory gaps simply do not exist in the first place.

However, in my earlier work I argued that this thought is too quick, and that in fact Finsler-Aczel and AFA face the very same kind of problem that Boffa faces. While it is true that Finsler-Aczel and AFA only countenance one self-singleton, so of course they never face the question of why two distinct self-singletons are distinct, I argued that they nevertheless face the question of why that self-singleton is that self-singleton, and I argued that this question admits of no good answer, and hence results in exactly the same kind of explanatory gap as Boffa is committed to. Here is what I said[8]:

> Both Finsler–Aczel set theory and AFA allow that there are self-singletons. They only differ from Boffa set theory (in this respect) in that they each allow only one self-singleton whereas Boffa set theory says that there are (very) many. But I think that as soon as you admit any self-singletons, you end up with the kind of explanatory bruteness we've been looking at. After all, suppose there's only one

[8] Cameron (2022, p76).

> self-singleton, Ω, as the stronger non-well-founded set theories maintain. Why is Ω=Ω? Well, what makes a set identical to a set? That their members are identical. The sole member of Ω is Ω. So Ω=Ω because Ω=Ω? That is no explanation, any more than that Ω1≠Ω2 because their members, namely Ω1 and Ω2, are distinct. In each case, there's nothing informative to say about *why* the self-singletons are identical or distinct, because there is never more to say about why sets are identical or distinct than that their members are identical or distinct, and self-singletons are identical to their members. And that is true whether there is exactly one self-singleton, or two, or proper class many. No matter how many self-singletons there are, if *this* self-singleton is identical to/distinct from *that* self-singleton, there is nothing informative to be said about *why* that is. That's just how things are . . . explanatorily brute identity is just as good or bad as explanatorily brute distinctness.

I concluded that Boffa, Finsler-Aczel and AFA are on a par: they each face an inevitable explanatory gap when it comes to the identity or distinctness of self-singletons. If Boffa is true there are more gaps around, because it countenances many self-singletons, and other exotica such as the sets with the infinitely descending membership chains, or the pairs of sets each of which is the singleton of the other, etc. But, I argued, that is just more of the same kind of problem, and merely more instances of a problem don't make for a worse theory.

Let's grant, for the sake of argument, that more explanatory gaps of the same kind don't make for a worse theory. In the next section I will argue that, even so, I was wrong to claim that Finsler-Aczel and AFA are committed to the same kind of explanatory gap that Boffa is committed to.

3 Two Ways of Individuating Sets

I argued in my earlier work that *every* non-well-founded set theory must countenance explanatory gaps, because there will be some claims concerning the identity or distinctness of sets that simply do not admit of explanation. Even considering the relatively restrictive set-theoretic universes of Finsler-Aczel or AFA, in which there is only a single self-singleton, Ω, it is still, I argued, explanatorily brute that Ω=Ω.

I now think this is wrong. My mistake was assuming that the only possible explanation one could give for the identity or distinctness of some sets goes via the identity or distinctness of the members of those sets. As I said, "there is never more to say about why sets are identical or distinct than that their members are identical or distinct"[9]. But while I still think that this is true given Boffa set theory, I now think that it is *not* true given Finsler-Aczel or AFA. The Finsler-Aczel or AFA set theorist has resources to explain the identity or distinctness of sets that go beyond the identity or distinctness of their members and, I will argue, this lets them explain without circularity the identity of any self-singleton with any self-singleton. It is precisely

[9] Cameron (2022, p76).

because these extra explanatory resources are not available to the Boffa theorist that they are inevitably committed to such explanatory gaps. So the puzzle really is, as it initially seemed, a puzzle specifically facing Boffa set theory (at least, restricting our attention to these three non-well-founded set theories).

The solution is going to be that the Finsler-Aczel or AFA set theorist, but not the Boffa set theorist, can appeal to the *structural* features of sets to explain facts about the identity or distinctness of sets in such a way that closes the explanatory gaps that would arise were they only able to appeal to facts concerning the identity or distinctness of their members. This is going to involve some appeal to metaphysical notions, such as the notion of essence and of *what it is* to be a certain thing, and so I will start by saying something about how I am using these terms.

I will be appealing to a non-modal notion of essence, in the tradition of Fine[10]. To say that it is essential to A that it is F is to say that part of what makes A *that very thing* and not some other thing is that it is F. The essentialist claim entails a modal claim: if A is essentially F then, necessarily, if A exists then A is F. But the modal claim does not entail the essentialist claim: it being necessary that A is F if A exists does not entail that A is essentially F. To use Fine's famous example, it is necessary both that Socrates' singleton has Socrates as a member (should they exist), and that Socrates be a member of Socrates' singleton (should they exist), but while it is essential to Socrates' singleton that it has Socrates as a member, it is not essential to Socrates that he be a member of Socrates' singleton. And that is because while it is true that part of what makes Socrates' singleton that very set is that it has the very thing Socrates as a member (if it had Plato as its sole member, by contrast, it would not be Socrates' singleton, it would be Plato's singleton), it is no part of what makes Socrates the very being he is that is a member of any set. (What makes Socrates the very thing he is might involve facts concerning his psychological features, e.g., but not that he stands in the *is a member of* relation to any set.)

I will also appeal to *what it is* claims. '*What it is* to be A is to be F', as I will understand it, says that there is nothing more or less to some thing's being A than that it is F. This entails the essentialist claim that A is essentially F, but it is not entailed by it. After all, many things are (we can suppose) essentially human, including both you and I, but it is not the case that what it is to be me is to be human, because you are human but are not me. For some things, there might be no true non-trivial *what it is* claim of this form. Consider, for example, the square roots of -1, i and -i. They are both, plausibly, essentially numbers, and essentially the square root of -1. But it is no conceptual mistake to say that there is no true and non-trivial claim of the form 'What it is to be i is to be F', precisely on the grounds that any informative feature you could put in place of 'F' would be a feature that i shares with -i.

[10] Fine (1994).

Following Agustin Rayo[11] I will assume that there is a tight connection between *what it is* claims and explanation. If you believe that 'What it is for ϕ is for Ψ' is true then, by your lights, it is unintelligible that it should be the case that ϕ but not the case that Ψ, and thus there is no demand for explanation for it being the case that Ψ beyond it being the case that ϕ. What it is to be water is to be H2O; given that, it is unintelligible that there should be water in the river without there being H2O in the river; given that, there is no demand for explanation for there being H2O in the river beyond explaining there being water in the river. If someone were to say 'I understand that there is water in the river. But why is there H2O?' they would be making a mistake, for *what it is* to be water is to be H2O.

Now consider two claims concerning the individuation of sets:

> <u>Extensional Individuation</u>: For any set, S, whose members are the Xs, *what it is* to be the very set S just is to have all and only the Xs as members.
>
> <u>Structural Individuation</u>: For any set, S, whose membership structure is a certain way, *what it is* to be the very set S just is to have that membership structure.

Extensional Individuation is familiar. Indeed, I think any set theorist who accepts an extensional set theory should accept it. It entails both that for any set, S, whose members are the Xs that it is essential to S that it has the Xs as its members, and that, necessarily, any set which has the Xs as its members is the very set S. Structural Individuation is perhaps less familiar. By membership structure I mean the purely structural features concerning how a set is built up.[12] A description of the membership structure of S will be a description that will not involve the identity of any particular sets in S's transitive closure, but will instead describe how each of those sets is built up from other sets that are themselves described in purely structural terms. So for example, we can describe the set depicted in Fig.1 by talking about the identity of its members: it is the set which has exactly one member, the singleton of the empty set, which itself is described in terms of the identity of its members: the empty set. But we could also describe that set in purely structural terms, in terms of how it is built up, without saying anything about the identity of any of the particular sets involved: it is the set that has one member, a set that itself has one member, a set that has no members.

Perhaps it is hard to hear these descriptions as different. But if so, I think that is just because we are so familiar with ZFC, in which the descriptions never come apart in interesting ways. That is because, in ZFC, every (pure) set is built up from the empty set, and it does not seem interestingly different if we describe that common building block as 'the empty set' or 'a set that has no members'. But when we consider non-well-founded set theories, we can see how describing the structural features of a set and describing its members can be interestingly

[11] Rayo (2013). Cf. Dorr (2016).

[12] Cf. Barwise and Moss (1991, p36-37) and Incurvati (2020, p194).

different. Consider two of Boffa's distinct self-singletons: Ω1 and Ω2. I think there's a clear sense in which these two distinct self-singletons, while having different members, have the same membership structure: the way in which those sets are built up is exactly the same, it's just they use different ingredients to get to different end results. If we try to describe Ω1 and Ω2, they can each be described as a set containing exactly one member, themself. That is a purely structural description, just telling you how the set is built up; it doesn't tell you what particular sets go into the building - it doesn't tell you whether 'themself' is Ω1 or Ω2, e.g. And so we can see that if Boffa set theory is true, the structural description of a set and the description of it in terms of its particular members, are interestingly different: the structural description is one that can be shared with distinct sets, whereas the description in terms of its particular members of course cannot be, given the principle of extensionality (which is, of course, much weaker than the principle Extensional Individuation). We will only get a description that applies to one and not the other of these self-singletons when we say something that goes beyond the purely structural features, i.e. something that invokes the identity of a particular set involved in the make-up of that set. So, e.g., Ω1 can be described as a set containing exactly Ω1 as a member, whereas Ω2 cannot be so described. This difference between them is not a difference in their membership structure: it's a difference not in how they are made up, but in what particular sets go into the make-up - a difference in the ingredients, not the recipe. Likewise, the sets (that exist according to Boffa) {Ω1, {Ω1,∅}, ∅} and {Ω2, {Ω2,∅}, ∅} are distinct but do not differ in their membership structure, for they are each built up in the same way from sets that do not differ in their membership structure: Ω1 is structurally identical to Ω2, and ∅ is structurally identical to ∅. Whereas even though they are built up from the very same sets, Ω1 and ∅, the sets {∅, Ω1} and {Ω1, {Ω1,∅}, ∅} are not structurally identical, because they are built up in different ways from those sets: same ingredients, but different recipe.

A set's membership structure is what is captured by a graph that exactly depicts it. The graph in Fig.1 does not itself say anything about the particular sets depicted by any given node. It does not tell you what particular set is depicted by node 1.3, only that it depicts a set that has no members. Now of course, we know that node 1.3 depicts a particular set - the empty set. But that is because we know that there is only one set that has no members. The graph doesn't tell us that, our set theory tells us that. Were someone to accept a non-extensional set theory that allowed multiple empty sets, the graph in Fig.1 would exactly depict each of the multiple sets that are a singleton of a singleton of an empty set. Those sets would be structurally identical, but distinct. To learn that a graph, g, exactly depicts a set, S, is to learn what S's membership structure is, and it is then a matter of what set theory we accept whether this by itself determines the particular members of S. If we accept ZFC, the answer is always that it does, because every set is built up in a unique way from a set whose membership structure - i.e. that it has no members - determines its particular identity - i.e. that it is the unique empty set.

However, if Boffa set theory is true, membership structure does not always determine the particular members of a set, because distinct sets (with, thereby, different members) can have the same membership structure, and that is shown by the fact that distinct sets can be exactly depicted by the same graph. Every one of the infinitely many self-singletons is exactly depicted by Fig.3; every one of the infinitely many sets that has as its sole member a distinct set that has its sole member yet another distinct set that has as its sole member yet another distinct set that has . . . and so on *ad infinitum* is exactly depicted by Fig.5; every one of the infinitely many sets that it is a singleton of a set that is itself a singleton of the original set is exactly depicted by Fig.7; etc. In the Boffa set-theoretic universe, membership structure massively underdetermines the particular members of a set, because infinitely many distinct sets, each with different particular sets as members, can share the same membership structure: they are built up in the same manner, but they built up in the same manner *from* different particular sets, resulting in they themselves being different sets.

As a result, the Boffa set theorist must reject Structural Individuation. It cannot be the case that *what it is* to be a particular set S just is to have S's membership structure, given that a different set could have the very same membership structure. It might be *essential* to any set in the Boffa set-theoretic universe that it has the membership structure it has, but it doesn't tell you *what it is* to be that very set, because having that membership structure is compatible with it being some other set.

Finsler-Aczel and AFA set theories, by contrast, can and should accept Structural Individuation. On both those theories, membership structure determines the identity of a set, because you never have a single graph exactly depicting multiple sets. Not only is membership structure essential to a set, on these views it is an individuating feature: once you know the membership structure of a set, you know what particular set it is, and that it is not some other set, because there is no other set that could have that particular membership structure: that is simply what it is to be that very set, that it be built up in that manner.

This does not mean that Finsler-Aczel and AFA have to deny Extensional Individuation. Both claims can be true. What it is to be the number 2 is to be the unique even prime, but also what it is to be the number 2 is to be the successor of 1. There can be multiple true and non-trivial *what it is* claims that give us different ways to grasp what makes a thing that very thing and not some other thing. And so, for some set S, whose members are the Xs, and that is exactly depicted by graph g, it might be true both that (as Extensional Individuation says) what it is to be S is to have the Xs as its members, and that (as Structural Individuation says) what it is to be S is to have the membership structure that is had by any set that is exactly depicted by g.

It is not forced on one who accepts Finsler-Aczel or AFA to accept Structural Individuation. Just because no two distinct sets share the same membership structure, even as a matter of necessity, does not entail that what it is to be a given set is to have its membership structure. After all, necessarily, no two people share the same singleton, but that does not mean (following Fine) that what it is to be a particular person is to belong to a particular singleton. It is no conceptual mistake for the defender of Finsler-Aczel or AFA to hold that it is an interesting modal truth that, necessarily, no two sets share the same membership structure, but that it is no part of what it is to be a particular set that it has the membership structure it has. But what's important for present purposes is that the defender of Finsler-Aczel or AFA *can* accept Structural Individuation, unlike the defender of Boffa set theory. And further, I argue that they *should* accept it, because doing so closes the explanatory gaps that otherwise would arise.

The shape of the response is likely by now predictable. If the defender of Finsler-Aczel or AFA does not accept Structural Individuation but only Extensional Individuation, then my earlier claim was correct that the only resources they have to explain the identity or distinctness of any sets is in terms of the identity or distinctness of their members. This does indeed yield explanatory gaps of the same kind that inevitably arise given Boffa set theory: there is no good explanation as to why this self-singleton is that self-singleton, for the only possible explanation is that they share their members, and the identity of their members is the very fact we set out to explain. But if Structural Individuation is true, we can explain the identity or distinctness of sets not just by pointing to the identity or distinctness of their members, but also by pointing to the similarity or difference of their membership structure. 'S1=S2 because they have the same membership structure' can be a good explanation. And of course, the sets having the same membership structure entails that they have the same members if Structural Individuation is true (assuming the principle of extensionality). But their having the same membership structure can *explain* their identity when having the same members cannot - when, for example, their being identical simply *is* their having the same members (as is the case when the set in question is a self-singleton).

So if Finsler-Aczel or AFA are true *and* Structural Individuation is true we can explain the identity of some self-singleton Ω1 with any self-singleton Ω2 as follows: they are identical because they have the same membership structure - that is to say, Ω1= Ω2 because they are each a set containing exactly one member, itself. Or if you think 'itself' smuggles in the particular identity of the set in question in an objectionable way, we can say: Ω1= Ω2 because they are each a set containing exactly one member that has exactly one member that has exactly one member . . . and so on. The only possible set that has that membership structure - the membership structure represented by the graph in Fig.3 - is the unique self-singleton, according to those set theories. This is a genuine explanation; the explanans is not the explanandum, for even if it is true that sets with the same membership structure are identical, that sets S1 and S2 have the same membership structure is clearly not the same fact as the fact that S1=S2. (After all, in Boffa, distinct sets can have the same membership structure, and even if Boffa is false, it is not conceptually incoherent.)

I conclude that I was wrong in my previous work to claim that Finsler-Aczel and AFA must countenance explanatory gaps of the same kind as Boffa: Boffa cannot explain the identity or distinctness of self-singletons, etc., but Finsler-Aczel and AFA *can* explain the identity of any self-singleton and any self-singleton, *provided* they accept Structural Individuation. Accepting Structural Individuation is, however, not an option for Boffa, as we have seen, since that theory posits many distinct sets with the same membership structure.

And the defender of Finsler-Aczel or AFA *should* accept Structural Individuation, I argue. One reason they should do is precisely so they can avoid these explanatory gaps. But also, while it is possible for the defender of Finsler-Aczel or AFA to follow the Boffa theorist in accepting Extensional Individuation but denying Structural Individuation, the resulting combination of views is somewhat strange, I think. Consider again the unique (given Finsler-Aczel or AFA) self-singleton, Ω. Even though all defenders of Finsler-Aczel or AFA must agree that there is only one possible set, Ω, that has itself as its sole member, a defender of Finsler-Aczel or AFA who denies Structural Individuation should allow that it is conceptually possible in some sense at least that there be a self-singleton that is not Ω. There is nothing conceptually incoherent - nothing that violates the very notion of a set - in the idea of a self-singleton that is not Ω, if what it is to be a particular set is merely to have certain particular members (and not to have a certain membership structure). The imagined self-singleton that is not Ω, call it 'Ω*', does not, after all, have the members that Ω has: for Ω* has Ω* as its sole member and Ω has Ω as its sole member, and Ω and Ω* are, ex hypothesi, not the same thing. So nothing about the nature of sets - assuming only Extensional Individuation - rules out the existence of multiple self-singletons. If Finsler-Aczel, or AFA, are true then there is, of necessity, only one self-singleton; but if, in addition, Structural Individuation is false, then nothing about what it is to be a set ensures that this is the case. That is odd, I think: it is as if there's a modal coincidence concerning the extent of the set-theoretic universe. It's almost like a haecceitist view where you hold that the qualitative features of an individual do not conceptually determine what particular individual exists, and yet there is a metaphysically necessary truth concerning what particular thing exists given that a thing with a particular qualitative profile exists. There's nothing inconsistent about such a view, but it leaves us wondering *why* some different individual with the same qualitative profile couldn't exist. By contrast, an anti-haecceitist has nothing to explain: if *what it is* to be a particular individual *just is* to have a certain qualitative profile then of course it is impossible that some thing exist and have the same qualitative profile as this thing, Frank let's say, without it thereby being the very thing Frank. The defender of Finsler-Aczel or AFA, if all they have to appeal to is Extensional Individuation, thinks that there being a self-singleton leaves it conceptually open what particular set exists, but it's like there's a single self-singleton haecceity, Ω-ness, that is the only one that can possibly be instantiated. Consistent, but unsatisfying. By contrast, a defender of Finsler-Aczel or AFA who accepts Structural Individuation can explain *why* there is no possible self-singleton other than Ω: because 'Ω' names a set with a certain membership structure, and *what it is* to be the very set Ω is to have that membership structure, so nothing could have that very membership structure - and so nothing could be a self-singleton - without being that very set Ω. It is conceptually incoherent for there to be a self-singleton that is not Ω; it is ruled out by the very nature of sets, by what it is to be a set. I think this gives the defender of Finsler-Aczel or AFA a reason to accept Structural Individuation; they can then explain a feature of the set-theoretic universe that goes unexplained if they accept Extensional Individuation only.

So contra what I said in my previous work, it is not true that all non-well-founded set theories inevitably yield these explanatory gaps, with Boffa just having more of them because it has more sets. It is non-well-foundedness without Structural Individuation that yields explanatory gaps.

Provided they accept Structural Individuation, Finsler-Aczel and AFA can close the explanatory gaps that they would otherwise have if they accepted Extensional Individuation only, because they have a way of explaining the identity of a set that doesn't involve explaining the identity of its members. And foundationalist theories like ZFC don't get into the situation of having these apparent gaps in the first place, since every set is built up from the same starting point - the empty set - whose nature we can get a grip on (and whether we think of its nature as determined by its particular members or its membership structure makes no difference: what it is to be that very set is to be a set that has no members).

This section ends on a pessimistic note for Boffa set theory, then, for it looks as though, contra what I argued previously, it is the worst of all the non-well-founded set theories we are considering, for it results in many explanatory gaps that can be avoided by its rivals. In the next section, however, I will argue that the Boffa set theorist need not be concerned.

4 Explanation, Arbitrariness, and Plenitude: a defensive move for Boffa

The goal of this section and the next is to defend Boffa against the objection to it we ended with in section 3 (this section), and to make a case for it over its rival non-well-founded set theories (section 5). (I will not be arguing for Boffa over ZFC. The issue of whether we should allow foundationless sets at all is left for another time; the goal here is simply to compare the anti-foundational options.)

To make this case I am going to rely on some analogies between the set-theoretic cases we have been looking at and some cases adjacent to the semantic paradoxes. I am going to argue that the best thing to say for the semantic paradoxes shows us that there is a good thing for the Boffa theorist to say in the face of the explanatory gaps that we have seen, and indeed that the analogies can show us that there are benefits to Boffa over Finsler-Aczel or AFA.

Let's start with the truth-teller sentence:

TT: TT is true

The puzzle with TT, of course, is that it could be true, or false, but nothing seems to determine which. We might be tempted to say, in that case, that it should be neither: but it can't be neither, because then it would not be true, in which case it would be false since it *says* that it is true. But it can consistently be true in which case it truly says that it is true, and it can consistently be false in which case it falsely says that it is true. But there seems to be no reason to favor one over the other.

Here's the response to this puzzle that I like.[13] There is not a single truth-teller sentence. In fact, there are infinitely many:

---

[13] See Cameron (2022, Ch.5).

TT1: TT1 is true
TT2: TT2 is true
TT3: TT3 is true
.
.
.
etc.

I think that some of these (indeed, infinitely many of them) are true, and some of them (indeed, infinitely many of them) are false. There is no explanation for why a true one is true, or why a false one is false. They just are. There is nothing outside of them that determines their truth-value. When we focus on a single truth-teller sentence it looks objectionably arbitrary to say that it is true, or that it is false, but nothing explains why it has the truth-value it has, and there is nothing outside of the sentence that determines its truth-value. But here is the slogan for this section: *plenitude dispels arbitrariness*. If there are infinitely many true truth-teller sentences, and infinitely many false ones, and you pick one and it is true, then while there is no explanation as to why it is true, there is nothing objectionably *arbitrary* about its having that particular truth-value rather than the alternative. There is no sense in which reality has privileged one option over another without a reason to do so: all options are realized, and you just so happen to have stumbled upon this one.

A truth-teller sentence is a little like a self-singleton: a self-referential entity, whose truth-value or identity is not determined by anything outside of itself, resulting in an inevitable lack of explanation as to why it has that particular truth-value, or why it is that particular self-singleton. But Boffa has very many self-singletons, and again: plenitude dispels arbitrariness. Consider again the combination of Finsler-Aczel (or AFA) with only Extensional Individuation and not Structural Individuation: on this view, as a matter of metaphysical necessity there is only one self-singleton, but because the nature of a set is determined solely by its members, and not by its structural features, it is conceptually possible that there be a different self-singleton. In that case, there is a puzzle as to why it is this particular self-singleton, Ω, that exists. Why not the conceptually possible distinct self-singleton Ω2, say? On this view there is an objectionable arbitrariness: while it is conceptually possible that there be other self-singletons, there is in fact, necessarily, only this one. Why that particular one? There is no explanation. Reality seems to have settled - arbitrarily - on one option, with no reason.

Of course, a defender of Finsler-Aczel shouldn't accept only Extensional Individuation, they should accept Structural Individuation, and say that there is no sense to be made of distinct self-singletons, for that would require there to be distinct sets with the same membership structure, which goes against the nature of sets. There is no arbitrariness on this view, because there are no multiple conceptually possible options for reality to choose between. But, I argue, Boffa dispels the arbitrariness by embracing a plenitude of self-singletons: there is no arbitrariness because reality selects *all* the options. Compare the multiverse solution to the fine-tuning puzzle[14]: there might be no *explanation* as to why the laws of nature are just as they need to be to allow for there to be life, but if every possible set of laws obtain at some part of the multiverse there is nothing objectionably *arbitrary* in the fact that those particular laws obtain in this particular corner of reality - everything obtains somewhere, and we just happen to be here. Or compare the argument often given for adopting a plenitudinous material ontology: if there is exactly one person-like object around in my vicinity that has some of its properties essentially and others accidentally, it might seem objectionably arbitrary that it has that particular modal profile. But if there is a plenitude of person-like objects around, and every possible modal profile is had by one of them, then while for any particular one there might be no good *explanation* to be had as to why it has that particular modal profile, it doesn't seem objectionably *arbitrary*: all the options are out there and you just happen to have stumbled upon this one.[15]

Likewise, pick one of the many self-singletons that Boffa believes in: Ωn. Why is it Ωn, and not Ωm? There is no explanation. But there is nothing arbitrary about this: both Ωn and Ωm exist, along with infinitely many other self-singletons. We just happened to land on Ωn. Similarly, let 'Ωr' be a self-singleton and 'Ωs' be a self-singleton. Is Ωr=Ωs? Nothing we've said tells us. Of course, Ωr=Ωs just in case they have the same members. But that is just to say that Ωr=Ωs iff Ωr=Ωs; that is no help. Of course, we could *stipulate* the answer. Let's stipulate that 'Ωr' and 'Ωs' pick out distinct self-singletons - a safe stipulation, since we know (given Boffa) that there are many such sets to choose from. Now we know that Ωr≠Ωs, but there is still no *explanation* to be had as to why Ωr≠Ωs, since the only explanation for the distinctness of sets is (if Boffa is true) via the distinctness of their members, and in this case the sets just are their members. But even though there is an inevitable explanatory gap here, there is nothing objectionably *arbitrary*: there is simply a plenitude of self-singletons, and we stipulated that we were choosing two distinct ones, so of course we end up with a self-singleton and a self-singleton that are not identical.

Let's look at another semantic case. Take the following sequence which is somewhat like the truth-teller unpacked into an infinite sequence:

P1: P3 is true
P3: P5 is true
P5: P7 is true
P7: P9 is true
.
.
.

---

[14] See, e.g., Friederich (2022, section 4).

[15] This argument for plenitude, while implicit in a lot of discussions, is most explicit in Bennett (2004). See also Fairchild (2019, 2022) for relevant discussion.

It is consistent for each sentence in this sequence to be true, and it is consistent for each sentence in the sequence to be false, but it is inconsistent for some to be true and some to be false. As with the solitary truth-teller, either option is available but, again like the truth-teller, nothing seems to determine which. There is nothing outside of the sequence to determine that the sentences are all true, or all false, since the sentences don't talk about anything other than the truth-value of other sentences in the sequence. Suppose they are all true. There is no explanation as to why they are all true as opposed to all false (likewise, mutatis mutandis, if they are all false): the explanation for the truth of any *individual* sentence on the list is that the very next sentence on the list is true, but those explanations of the truth-value of the individual sentences *presupposes* that we have a list of true sentences as opposed to false ones. Nothing explains why we have a list of truths in the first place as opposed to a list of falsehoods.

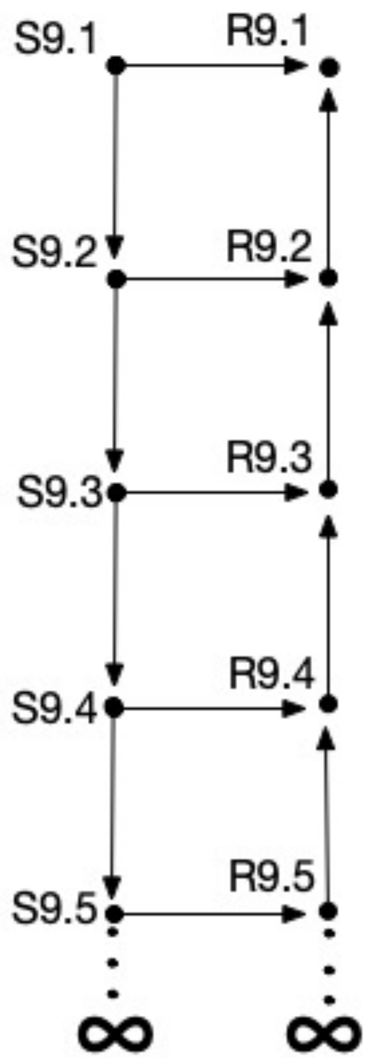
If there was but one such infinite sequence of sentences, whatever truth-value those sentences had would be entirely arbitrary: if they were all true, e.g., it would seem that reality is arbitrarily privileging truth, for no reason. However, there are infinitely many such sequences - and plenitude dispels arbitrariness. As well as the sentences P1, P3, P5, etc., there is

P2: P4 is true
P4: P6 is true
P6: P8 is true
P8: P10 is true
.
.
.

. . . and infinitely many such sequences. Some of them (infinitely many) will consist entirely of truths, some of them (infinitely many) entirely of falsehoods. Now pick one. Perhaps it consists entirely of truths. There is nothing to explain why *these* particular sentences are all true; but there is nothing *arbitrary* about them being so - there are sequences of truths, and there are sequences of falsehoods, and we happen to have landed on a sequence of truths. If reality is complete with every option, there is nothing arbitrary in discovering any particular option.

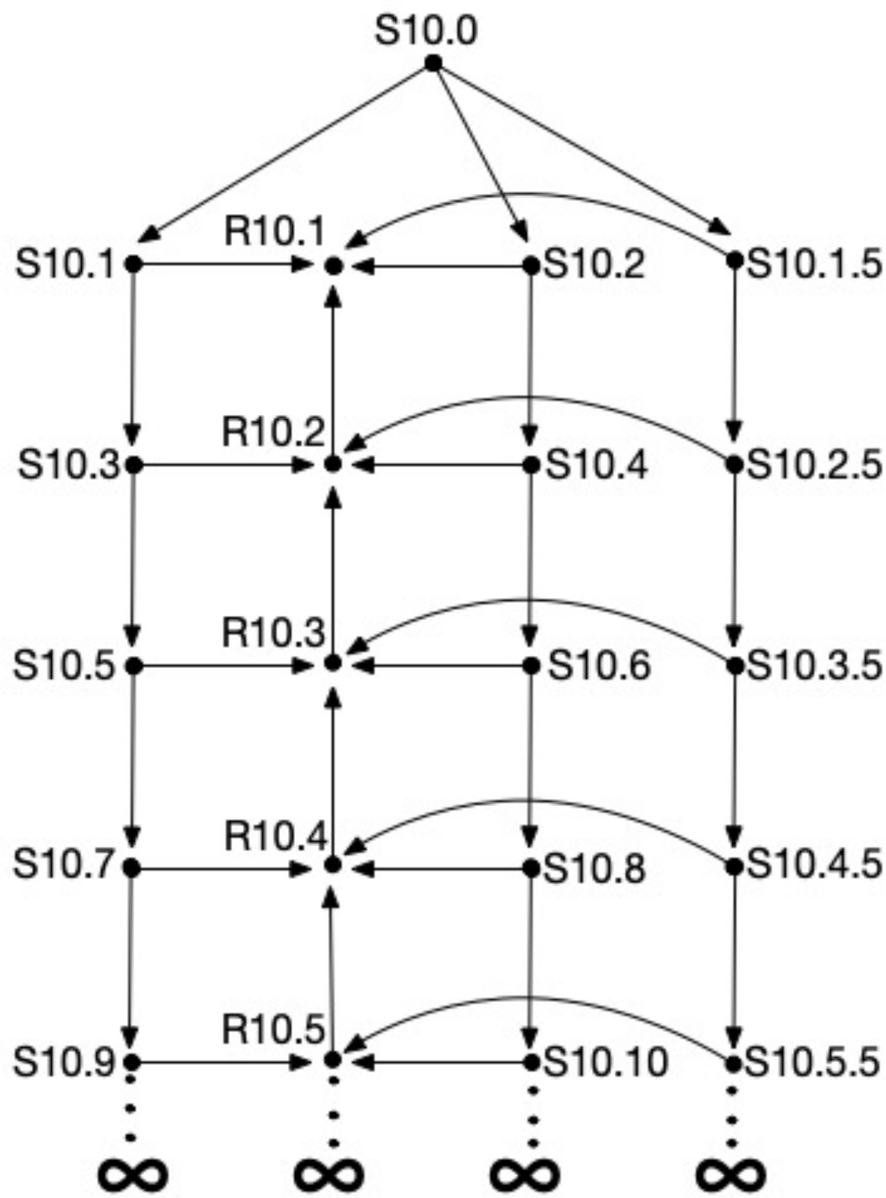
Compare that to the pair of infinitely descending sequences of sets: S1={S3}={{S5}} . . . , S2={S4}={{S6}}. . . These are two of infinitely many such sequences, if Boffa is correct. And as we saw above, we're left with an inevitable explanatory gap as to why S1 is that very set and not the set S2 (etc.). But again, the very fact that there is a plenitude of such sequences dispels, I think, any sense of arbitrariness. It would be objectionably arbitrary if there were just one such sequence and we couldn't explain why it was made up of these particular sets. For example, if (*per impossibile*) the sequence starting with S1 existed but no other, it would be objectionably arbitrary *why* it was this sequence of sets that existed as opposed to, say, the sequence starting with S2. But plenitude dispels arbitrariness: every such sequence you can conceive of exists, given Boffa, so pick one, and while there might be no explanation to be had as to why it is that sequence of sets and not some other, there is nothing objectionably arbitrary about it being that sequence: that's just the one you've stumbled upon.

In this case, we cannot use Finsler-Aczel or AFA with Extensional Individuation and not Structural Individuation to illustrate the badness of having just one such sequence while holding that the nature of sets is determined solely by their members and not their structural features, because Finsler-Aczel doesn't believe in *any* such sequence of sets, where each set has the next distinct set on the infinite chain as its sole member. However, a closely related case will illustrate the point. Consider the graph in Fig.9.



The sequence of sets S9.1, S9.2, S9.3, . . . is an infinite sequence of sets the first of which contains the second set in the sequence and the empty set, the second of which contains the third set in the sequence and the singleton of the empty set, the third of which contains the fourth set in the sequence and the singleton of the singleton of the empty set . . . and so on, *ad infinitum*.

This sequence of sets exists in each of Boffa, Finsler-Aczel, and AFA. Where Boffa differs from Finsler-Aczel and AFA - and this will be no surprise at this point - is that while Finsler-Aczel and AFA only allow there to be one such infinite sequence of sets, Boffa says that there are infinitely many. Consider the graph in Fig.10



Each of the ’S-sequence’ paths beneath S10.0 are isomorphic to one another, so this graph is not an exact depiction of a set according to Finsler-Aczel or AFA. But it is according to Boffa, since no two nodes share the same children: and of course we could add as many further paths to this tree as we like with the same structure. So in Finsler-Aczel and AFA there is exactly one set that has this structure, but in Boffa there are infinitely many.

So now consider Finsler-Aczel (or AFA) with only Extensional Individuation and not Structural Individuation. There is exactly one set (represented by node S9.1 on Fig.9) that has two members: the empty set and a set that has two members: the singleton of the empty set and a set that has two members: the singleton of the singleton of the empty set and a set that has two members . . . and so on, ad infinitum. But which set is it? Is it the same as the set represented by node S10.1 on Fig.10 that Boffa believes in, or is it the same as the set represented by node S10.2, or S10.1.5, or . . . ? There are infinitely many sets like that according to Boffa set theory, and Finsler-Aczel and AFA believes in exactly one; but if sets are individuated only by their members and not their structural features then we can make conceptual sense, as least, of the multiple sets with this structure that Boffa believes in, so it seems like there should be a sensible question as to which of Boffa’s infinitely many sets of this structure is the unique one that Finsler-Aczel and AFA believes in. But there is nothing at all we can say to explain why it would be one of them rather than any other, so it would seem entirely arbitrary that it is this one that exists rather than some other. Of course, if the defender of Finsler-Aczel (or AFA) accepts Structural Individuation, they face no such arbitrariness: they will think that the defender of Boffa is not merely being ontologically profligate in accepting multiple sets with this structure, but rather they are making a conceptual mistake - there cannot be multiple sets with the same structure, because what it is to be a particular set is just to have a certain membership structure. Boffa, of course, has to deny Structural Individuation and so, as is now familiar, faces an inevitable explanatory gap. Nothing explains why the set that is represented by S10.1 on Fig.10 is that very set as opposed to, say, the set that is represented by S10.2, because the only difference between the sets is that the second member of each (the one that is not the empty set) is distinct - but those distinct members are themselves structurally identical, and so we just get the very same question again: why is this one this one and not that one? So there is an inevitable explanatory gap; but, once again, plentitude dispels arbitrariness - even though there is no *explanation* to be had as to why any of these sets with this structure is that particular set as opposed to any other set with the same structure, because there is an infinite plenitude of such sets, every conceptually possible option exists, and so there is no sense of reality arbitrarily admitting some but not another option without a reason.

Explanatory gaps are not by themselves objectionable, I claim. We should not expect that everything can be explained. But what is objectionable, I hold, is when the lack of explanation results in arbitrariness: when there is a sense of reality arbitrarily picking one option out of many, with no good explanation as to why. Perhaps such arbitrariness is unavoidable, but it is better to avoid it if possible. What I have tried to show is that in the case of non-well-founded set theories, there are two ways to avoid it. One way - open to defenders of Finsler-Aczel or AFA but not defenders of Boffa - is to accept Structural Individuation. In that case, there is no arbitrariness as to why the particular sole self-singleton that exists, say, is the one that exists, because we can explain why this is the only self-singleton that *could* exist: any self-singleton would, by its nature, be it, since what it is to be that very set just is to have the very membership structure of a self-singleton. Another way to avoid the arbitrariness, however - open to defenders of Boffa, but not Finsler-Aczel or AFA - is to embrace a plenitudinous set theory in which there are many explanatory gaps, but even though we cannot explain why a particular self-singleton, say, is that self-singleton and not some other, there is no sense in which it is arbitrary that that particular self-singleton exists, because *they all do*. Reality hasn't arbitrarily selected one option, it's let all the infinitely many flowers bloom. Plenitude dispels arbitrariness.

Nothing in this section gives us an argument for Boffa set theory, of course. This has all been defensive: to resist what looked like an argument *against* Boffa set theory. In the next section I shall make a positive case for Boffa (at least, among non-well-founded set theories).

5 Symmetry and Structure: an argument for Boffa

The argument for Boffa will go via arguing against Structural Individuation. Since a non-plenitudinous foundationless set theory without Structural Individuation results in objectionable arbitrariness - as we saw in section 4 - if we have an argument against Structural Individuation we have an argument for accepting either a foundationalist set theory like ZFC, or a plenitudinous non-well-founded set theory like Boffa. The issue of whether to accept a non-well-founded set theory in the first place is beyond the scope of this paper, so I will simply conclude that the conditional is true: if we are to accept a non-well-founded set theory, we should accept a plenitudinous one like Boffa, over a non-plenitudinous one like Finsler-Aczel or AFA.[16]

We will work up to the argument against Structural Individuation by first once again considering some analogies to the semantic paradoxes and related phenomena. In section 4, I defended (or at least asserted) the claim that there are multiple truth-teller sentences, some of which are true and some of which are false, as well as the claim that there are multiple infinite sequences of the 'unpacked truth-teller', some sequences consisting entirely of truths, others consisting entirely of falsehoods. Accepting this view commits one to the claim that there can be structurally identical sentences that differ in truth-value. Some philosophers have found that result untenable. Graham Priest, for example, appeals to the oddness of this commitment in his argument for a dialethic solution to the so-called no-no paradox. This consists of two sentences, each of which says the other is false:

NN1: NN2 is false
NN2: NN1 is false

---

[16] Obviously, this argument does not tell uniquely in favor of Boffa. It leaves on the table, for example, Quine's (1937) New Foundations. I'm only arguing for Boffa *given* the options I've (somewhat arbitrarily) put on the table in this paper: one has to limit one's attention somehow!

The paradox is that we can consistently assign one of these truth and the other falsity, but there seems to be nothing to determine which. I think we should simply embrace the conclusion that one is true and other false, but nothing *explains* why the true one is true and the other false, for nothing is relevant to the truth-value of any one other than the truth-value of the other one in the pair. But Priest objects to the idea that the 'symmetrical' sentences can have different truth-values. He says[17]:

> Consider a business card, on each side of which is written a single sentence: *the sentence on the other side is false*. . . [T]here is only one consistent assignment of truth values to these two sentences: one is true; the other false. But which is the true one? Since the situation is completely symmetrical, there is nothing in virtue of which it is one rather than the other. Despite this, classical logic assures us that it is one or the other. . . [T]he possibility that the two sentences in question have different semantic properties flies in the face of the truth-maker principle, that all truths have truth makers; but . . . [w]hat is much worse is that the possibility is a manifest *a priori* repugnance. The situation concerning the card is, in all respects, symmetrical; it cannot, therefore, have an asymmetric upshot. Either both sentences are true, or both are false. But if both sentences are true, they are false as well; and if both sentences are false, they are true as well. Hence, it would seem, both sentences are true and false.

Priest does not specify exactly what this symmetry consideration amounts to, but my best reconstruction is that he believes something like the following claim:

<u>Symmetry</u>: If Sx and Sy are 'structurally identical' then they have the same truth-value.

Symmetry is somewhat analogous to Structural Individuation: the former rules out structurally identical sentences with different truth-value, the latter rules out structurally identical sets with different identities.

I don't think we should accept Symmetry. And in seeing why Symmetry is false, I think we can also see why Structural Individuation is false. Consider the following four lists of sentences:

List 1:
At least two sentences on this list are F
2+2=4
2+3=5

List 2:
At least two sentences on this list are F
2+2=6
2+3=10

List 3:
At least one sentence on this list is T

List 4:
At least one sentence on this list is T

[17] Priest (2005, p689-690).

Nobody would argue, I take it, that the first sentence on List 1 and the first sentence on List 2 cannot differ in truth-value. They do differ in truth-value, because of the different truth-values of the other sentences on the list. However, in any sense of 'structurally identical' I can make sense of, they are structurally identical. So Symmetry cannot be true. And so why should we rule out the first (and only) sentences of List 3 and List 4 (which are basically just versions of the truth-teller) having different truth-values?

If anything like Symmetry is true, it must have a clause to exclude a difference in truth-value as a result of what's going on 'outside' of those sentences. Something like:

> Symmetry*: If Sx and Sy are 'structurally identical' then they can only have the same truth-value if there is something outside of those sentences that explains the difference in truth-value.

But I don't think there is good reason to accept Symmetry* given the falsity of Symmetry. If two sentences can be structurally identical and differ in truth-value as a result of what is happening in the world outside of them, why can't there be structurally identical sentences that differ in truth-value as a result of what is happening 'internal' to them: for example, just *because* one is true and the other false? That is, if it is okay to say that sentence 1 on List 1 is false because it is accompanied by truths, but the structurally identical sentence 1 on List 2 is true because it is accompanied by falsehoods, why is it not okay to say that sentence 1 on List 3 is true (e.g.) just because it is true and is on that list, whereas sentence 1 on List 4 is false (e.g.) just because it is false and is the only sentence on the list? Of course, we can't *explain* why one is true and the other false. But why think we should be able to explain everything? And we might have found it objectionably arbitrary to say that this one is true, or that that one is false, but we should no longer think so now that we know that plenitude dispels arbitrariness: there are infinitely many lists like lists 3 and 4: infinitely many of them consist of a single truth, infinitely many of a single falsehood. It is not arbitrary if you happen to have settled your attention on a true one (e.g.), it is just the luck of the draw.

I think what's really motivating the symmetry considerations is something like the following thought:

> Anchoring: The truth-value of any sentence must be explained by goings on in the non-semantic world.

The idea behind Anchoring is that semantic values don't come for free: there is truth, or falsity, only when a sentence 'meets the world' and correctly, or incorrectly, represents it. If Anchoring is true, it is fine for the first sentences of Lists 1 and 2 to differ in truth-value, despite being structurally identical, because they meet different parts of the world, and the worldly difference accounts for the semantic difference. But if Anchoring is true not only would we not be happy with the first sentences of Lists 3 and 4 *differing* in truth-value, we would not be happy with them having any truth-value, for they do not meet the world at all. If Anchoring is true, purely self-referential sentences like the truth-teller, or indeed the liar, do not have truth-values.

I think Priest's symmetry considerations are really a red herring, then. The fundamental idea here behind the objection to (inter alia) there being multiple truth-teller sentences with different truth-values is Anchoring: it is an objection to any of them having a truth-value in the first place.

Anchoring is a tempting thought but is ultimately, to my mind, unsatisfying. The best worked out view that starts from an Anchoring intuition, I think, is Tim Maudlin's[18]; but I think the view should be rejected due to a problem emphasized by Hartry Field.[19] I don't have the space to go into either Maudlin's view or Field's objection in detail here, but here's the brief gist. Maudlin thinks that atomic sentences that utilize non-semantic predicates get to be true or false as a result of meeting the world and correctly or incorrectly describing it, and any other sentence that is true or false gets to be so by being built up from those atomic sentences (for example, a conjunction of them will be true iff both the atomic sentences that are conjuncts are true, etc.). Sentences like the truth-teller or liar are defective - 'ungrounded' says Maudlin, following Kripke[20] - and don't get to have a truth-value. But then consider the liar, L: L is not true. Let's make this formal by utilizing a truth-predicate and naming the sentence: L: $\neg T(<L>)$. This is neither true nor false, on Maudlin's view. Views that say that face a familiar problem, of course: if it's neither true nor false, then it is not true: i.e. $\neg T(<\neg T(<L>)>)$. But L *just is* $\neg T(<L>)$, so saying that the liar is not true is the same as saying $\neg T(<L>)$, which is just to say the liar sentence. And so Maudlin's theory says both $\neg T(<L>)$ and $\neg T(<\neg T(<L>)>)$: that is, it says something while at the same time saying that this something is not true.[21] So the theory says of itself that parts of it are not true. Maudlin acknowledges and embraces this consequence, and of course there is lots more to be said, but I think this is a good reason to reject the theory. And I take this to be a reason to reject Anchoring: the guiding thought behind it sounds initially plausible, but just doesn't work out.

And so I think we should simply accept that the first sentences of lists 3 and 4 have truth-values. They don't need to meet the world beyond themselves to do so because Anchoring is false. And since the only thing that is relevant to their truth-value is their own truth-value, they might end up with different truth-values, and so Structure* (*a fortiori*, Structure) is false. Two sentences can be structurally identical, but one be true and one be false, even if there is nothing in the world beyond those sentences to account for the difference in truth-value.

Let's bring this back to set theory. In the semantic case, we saw that the naive symmetry condition (Symmetry) *must* be false, because distinct but structurally identical sentences can

---

[18] Maudlin (2004).

[19] Field (2006).

[20] Kripke (1975).

[21] Field (2006, p715).

meet different portions of the world, and this leads to a difference in truth-value. The analogous thought in the set-theoretic case is this: however plausible it might be that *pure* sets with the same membership structure are the very same set, this is obviously not plausible when we also take into account *impure* sets. {{},{a},{{a}}} and {{},{b},{{b}}} have the same membership structure: they are built up in exactly the same way from an urelement. But, provided that a≠b, they are distinct sets. Nobody, of course, finds that suspicious: the difference in *starting point* leads to a difference in what set you end up with, even if the sets are built up from that starting point in exactly the same way. And so just as in the semantic case, we might try to restrict our structuralist claim to a particular class of cases. In the semantic case, we retreated from Symmetry to Symmetry*: the structuralist intuition was meant to hold just for 'pure' sentences that don't meet the world, it wasn't meant to hold (so the thought goes) when the different truth-values of structurally identical sentences could be explained by those sentences being responsive to different portions of the non-semantic world. Analogously, in the set-theoretic case, we might restrict Structural Individuation to pure sets: the structuralist intuition is meant to hold just for pure sets that don't meet the world (i.e. that don't involve any non-sets in how they are built up), it is not meant to hold when the different identities of structurally identical sets can be explained by those sets being built up from different portions of the non-set-theoretic world (i.e. by them being built up from different urelemente).

Following the analogy, I think what's really behind the intuition for Structural Individuation is something like an anchoring intuition: you can have distinct *impure* sets with the same membership structure precisely because those sets 'meet' different portions of the world - the difference in identity of the sets is explained by the difference in identity of the urelemente from which they are built. But you can't have a 'merely set-theoretic' difference: you can't have distinct pure sets with the same membership structure, precisely because there is nothing beyond the set-theoretic realm to account for the difference in identity.

In the semantic case, I argued that once you allow different structurally identical sentences to differ in truth-value because they meet different portions of the world, you should also allow different structurally identical sentences to differ in truth-value just because there's a plenitude of such sentences, some are true and some are false, and you happen to have settled your attention on two that differ. Likewise, I think in the set-theoretic case that once you allow there to be sets with the same membership structure that are distinct as a result of being built up from different portions of the non-set-theoretic world, you should also allow there to be sets with the same membership structure that are distinct as a result of being built up from different portions of the set-theoretic world. It is no surprise that there can be distinct *impure* sets with the same membership structure, precisely because there are distinct urelemente around to build up distinct but structurally identical impure sets from. Similarly, then, it would be no surprise if there could be distinct *pure* sets with the same membership structure, *if* there are distinct set-theoretic starting points to build up such distinct pure sets from.

And that is exactly what Boffa gives us: it gives us a plenitude of such starting points, in the purely set-theoretic realm. Distinct self-singletons - such as Ω1 and Ω2 - might be structurally identical, but they are built up from different starting points just as the similarly structurally identical {a} and {b} are. Of course, in the case of Ω1 and Ω2, the starting points *just are* Ω1 and Ω2, but why should that make a difference? And we cannot *explain* the difference between those starting points - nothing explains why Ω1 is Ω1 and not Ω2, e.g. - but there's no guarantee that we can explain the difference between a and b either. a and b might be qualitatively indiscernible electrons. Unless you believe the identity of indiscernibles, you're going to think that there are (at least possibly) some non-sets that are just distinct with nothing illuminating to say as to *why* they are distinct. If so, it is no surprise and no objection to a set theory that it can construct distinct structurally identical impure sets from distinct non-sets whose distinctness cannot be explained. Why then should it be an objection to Boffa that it can construct distinct structurally identical pure sets from distinct sets whose distinctness cannot be explained? Likewise with the infinitely descending membership chains. In this case, there is no starting point, because each set in the construction of each of these sets is itself constructed from some further set; however, we can view each set on the chain as being constructed from an infinite sequence of sets. So each set is built from an infinite sequence of sets, and each of those infinite sequences is structurally identical. But if the multiple starting points we get from each self-singleton are unobjectionable, so should the multiple structurally identical infinite sequences of sets be. Again, we cannot *explain* the distinctness of one infinite sequence from another, but that is not by itself objectionable, I have argued.

The final claim I want to argue for is that we *ought* to accept Boffa's plenitude of set-theoretic starting points and resulting plenitude of distinct structurally identical pure sets. The reason for this is that it would be *ad hoc* to treat pure sets differently from impure sets in this respect unless there is a compelling reason to do so. But I think that there is no such reason; the explanatory gaps that result from allowing such a plenitude of structurally identical pure sets do not, I have argued, give us any such reason, precisely because the very plenitudinous nature of the resulting ontology dispels any otherwise objectionable arbitrariness that might have otherwise resulted from such a lack of explanation. The intuition behind treating pure and impure sets differently in this manner, I suggest, relies on a set-theoretic analogue of the Anchoring thought: that a difference in the identity of structurally identical sets must be accounted for by them meeting different portions of the world. {a} and {b} can be distinct, e.g., because one is built up from a and the other from b and (let us stipulate) a≠b. But this thought only gets us a difference between pure and impure sets if we demand that by 'meeting the world' we mean the non-set-theoretic world. After all, Ω1 and Ω2 also meet different portions of the world: Ω1 is built up from Ω1 and Ω2 is built up from Ω2, and Ω1≠Ω2. But if we can have distinct sets with the same membership structure because they meet different parts of the non-set-theoretic world, I think we should also allow that we can have distinct sets with the same membership structure because they meet different parts of the set-theoretic world. Just as in the semantic case we should accept that there can be differences in truth-value that are not explained by anything beyond some sentences being true and others false, with no explanation to be had as to why those particular sentences have those particular truth-values, so in the set-theoretic case we should accept that there can be differences in the identity of sets that are not explained by anything beyond this set being this one and that set being that one, with no explanation to be had as to why a particular set is that very set and not some other.

References

Aczel, Peter, (1988), *Non-well-founded Sets,* Stanford: CSLI. https://doi.org/10.2307/2274782

Barwise, Jon and Moss, Larry, (1991), 'Hypersets', *The Mathematical Intelligencer* 13(4), p31-41 https://doi.org/10.1007/BF03028340

Bennett, Karen, (2004), 'Spatio-temporal Coincidence and the Grounding Problem', *Philosophical Studies* 118, p339-371 https://doi.org/10.1023/B:PHIL.0000026471.20355.54

Black, Max, (1952), 'The Identity of Indiscernibles', *Mind* 61, p153-164 https://doi.org/10.1093/mind/lxi.242.153

Bliss, Ricki Leigh, (2013), 'Viciousness and the Structure of Reality', *Philosophical Studies* 166(2), p399-418 https://doi.org/10.1007/s11098-012-0043-0

Boffa, Maurice, (1969), 'Sur la Theorie des Ensembles sans Axiome de Fondement', *Bulletin de la Société Mathématique de Belgique* 21, p16-56 https://doi.org/10.1002/malq.19720180407

Cameron, Ross, (2022), *Chains of Being*: *Infinite Regress, Circularity, and Metaphysical Explanation*, Oxford: Oxford University Press https://doi.org/10.1093/oso/9780198854272.001.0001

Dorr, Cian, (2016), 'To Be F Is To Be G', *Philosophical Perspectives* 30, p39-134 https://doi.org/10.1111/phpe.12079

Fairchild, Maegan, (2019), 'The Barest Flutter of the Smallest Leaf: Understanding Material Plenitude', *The Philosophical Review* 128(2), p143-178 https://doi.org/10.1215/00318108-7374932

Fairchild, Maegan, (2022), 'Arbitrariness and the Long Road to Permissivism', *Nous* 56(3), p619-638 https://doi.org/10.1111/nous.12376

Field, Hartry, (2006), 'Maudlin's Truth and Paradox', *Philosophy and Phenomenological Research* 73(3), p713-720 https://doi.org/10.1111/j.1933-1592.2006.tb00557.x

Fine, Kit, (1994), 'Essence and Modality', *Philosophical Perspectives* 8, p1-16 https://doi.org/10.2307/2214160

Friederich, Simon, (2022), 'Fine-Tuning', The Stanford Encyclopedia of Philosophy (Summer 2022 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2022/entries/fine-tuning/>

Incurvati, Luca, (2014), 'The Graph Conception of Set', *Journal of Philosophical Logic* 43(1), p181-208 https://doi.org/10.1007/s10992-012-9259-x

Incurvati, Luca, (2020), *Conceptions of Set and the Foundations of Mathematics*, Cambridge: Cambridge University Press https://doi.org/10.1017/9781108596961

Kripke, Saul, (1975), 'Outline of a Theory of Truth', *Journal of Philosophy* 72(19), p690-716 https://doi.org/10.2307/2024634

Maudlin, Tim, (2004), *Truth and Paradox: Solving the Riddles*, Oxford: Oxford University Press https://doi.org/10.1215/00318108-2007-010

Priest, Graham, (2005), 'Words Without Knowledge', *Philosophy and Phenomenological Research* 71(3), p686-694 https://doi.org/10.1111/j.1933-1592.2005.tb00481.x

Priest, Graham, (2014), *One*, Oxford: Oxford University Press https://doi.org/10.1093/acprof:oso/9780199688258.001.0001
Quine, W.V., (1937), 'New Foundations for Mathematical Logic', *American Mathematical Monthly* 44, reprinted in his *From a Logical Point of View*, Harvard University Press, (1953), p85-101 https://doi.org/10.2307/2267377
Rayo, Agustin, (2013), *The Construction of Logical Space*, Oxford: Oxford University Press https://doi.org/10.1093/acprof:oso/9780199662623.001.0001
Rieger, Adam, (2000), 'An Argument for Finsler-Aczel Set Theory', *Mind* 109(434), p241–254 https://doi.org/10.1093/mind/109.434.241