THE REVIEW OF SYMBOLIC LOGIC
Volume 1, Number 1, June 2008

THE ITERATIVE CONCEPTION OF SET

THOMAS FORSTER
Centre for Mathematical Sciences

Abstract. The two expressions 'The cumulative hierarchy' and 'The iterative conception of sets' are usually taken to be synonymous. However, the second is more general than the first, in that there are recursive procedures that generate some ill-founded sets in addition to well-founded sets. The interesting question is whether or not the arguments in favour of the more restrictive version – the cumulative hierarchy – were all along arguments for the more general version.

The phrase 'The iterative conception of sets' conjures up a picture of a particular set-theoretic universe – the cumulative hierarchy – and the constant conjunction of phrase-with-picture is so reliable that people tend to think that the cumulative hierarchy is all there is to the iterative conception of sets: if you conceive sets iteratively, then the result is the cumulative hierarchy. In this paper, I shall be arguing that this is a mistake: the iterative conception of set is a good one, for all the usual reasons. However, the cumulative hierarchy is merely one way among many of working out this conception, and arguments in favour of an iterative conception have been mistaken for arguments in favour of this one special instance of it. (This may be the point to get out of the way the observation that although philosophers of mathematics write of the *iterative* conception of set, what they really mean – in the terminology of modern computer science at least – is the *recursive* conception of sets. Nevertheless, having got that quibble off my chest, I shall continue to write of the *iterative conception* like everyone else.)

1. The cumulative hierarchy. There is a celebrated observation of Quine's (1968) "No entity without identity," which throws down a challenge to theory designers everywhere. If you lack a satisfactory identity criterion for widgets, then you cannot use first-order predicate calculus with equality to theorise about them; that is to say, you are unable to treat them formally. In particular, any story about what sets *are* had better include a chapter in which we learn how to tell when two sets are the same and when they are different.

The cumulative hierarchy gives an entirely satisfactory response to this challenge; two sets are identical if every member of the one is identical to a member of the other and *vice versa*. Since $\in$ is well-founded in the cumulative hierarchy, this regress must terminate; then, the fact that it terminates gives us an unequivocal and intelligible criterion for identity between sets.

I argued in Forster (1983) that it is precisely this feature of the cumulative hierarchy that makes it so attractive. I do not know who was the first person to make this point: when making it in Forster (1983), I assumed I was merely giving routine expression to an uncontroversial common understanding. Certainly one of the – uncontroversial – points

Received: October 25, 2007; in revised form: February 1, 2008

© 2008 Association for Symbolic Logic
doi:10.1017/S1755020308080064

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

routinely made in arguing against accepting self-membered sets is that such an acceptance sabotages the possibility of a recursive algorithm for deciding identity.

I now think that this is an idea worth taking further, and accordingly I am taking it as my point of departure here.

> The appeal of the cumulative hierarchy lies precisely in its neat response to Quine's challenge. (*Q*)

If (Q) is correct, then any concept of set that has a response to Quine's challenge that is as good as that of the cumulative hierarchy has an equally good a claim to the title of *the* way to think about sets. There are other iterative ways of generating sets, and the iterative nature of their genesis will guarantee a similarly satisfactory recursive account of equality between the sets so generated. Some of these ways generate more than just the well-founded sets, so my conclusion will be that the iterative conception of sets will give us a great deal more than just the well-founded sets.

Let us start off with a rehearsal of the cumulative hierarchy conception with a view to isolating some key features. I shall be exploiting the image (which Boolos, 1971, attributes to Kripke) of the lasso. In the cumulative hierarchy story, we lasso collections of sets, and then – before throwing them back into the herd of sets whence we plucked them – we perform some magic on the lasso contents (otherwise we would not get a new set). The magic is performed with the aid of a *wand.*

(There are various points that can be made about how the things one catches with one's lasso are not earlier sets, but *copies* of earlier sets, since a set can get caught in a lasso with other sets more than once. Perhaps we mean *token*s of earlier sets. A rigorous description of this process in terms of lassos probably does need to exploit the type–token distinction or something very like it. But we probably do not need to worry about such subtleties – at least for the moment.)

The point I want to emphasise here is the indispensability of the wand. The point needs making because if the only iteratively conceived sets one examines are the sets of the cumulative hierarchy, then it is possible for a hurried thinker to overlook the need for the wand, and this is a mistake that leads to further error. We need the wand for the following reasons (among others).

1. There is first the apparently banal point that I have just made that unless one does something to the lasso contents before throwing them back, then one has not done anything to increase the stock of sets: one has merely picked some up and put them back. The function performed by the wand is the turning of a preset into a set. However, the point can be made in another way.
2. As (Q) emphasises, one very pleasing aspect of the cumulative hierarchy picture is the consistent and fitting story it tells about the nature of identity between sets. We know from the axiom of extensionality that two sets are identical iff their members are identical. This sounds as if we can decide all questions like $x = y$? in the same way: just check whether every member of x is a member of y and *vice versa*. However, if that were all there was to it, then we would be in a situation where *any* question '$u = v$?' could be answered *immediately*, and the extra information provided by the news that all sets belong to the cumulative hierarchy would not enable us to do anything we could not do already. Life was never meant to be that easy. If the assumption that every set is in the cumulative hierarchy is to be any use to us, we will have to restrict ourselves to supposing that we can only (i) tell whether or not a set is empty and (ii) recover the members of a set on demand,

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

with all questions about those members to be answered later. The insight that best underpins (ii) is the idea that a set is magicked out of lasso contents and that the magicking can be reversed and the contents recovered.

Once we are clear in our minds that sets need a wand as well as a lasso, we can make the cumulative hierarchy work for us, as follows. To determine whether or not $x = y$, we take x and y apart and check, for each member of x whether or not there is a member of y to which it is identical (and *vice versa* too of course). By reducing the question of identity between x and y to identity between members of x and members of y, we obtain a recursive algorithm for testing equality.

Where does the process of the last paragraph bottom out? With the empty set: the one time when we can give an *immediate* answer to the question 'Is $x = y$?' is when one or other of x and y is empty. If they are both empty, the answer is 'yes'; if only one of them is empty, it is 'no'. Must we always reach such a stage? Yes, because at each stage in this recursion, we are asking questions about sets of ever-decreasing rank, and the ordinals are well-founded: every descending sequence of ordinals is finite.

We can characterise this algorithm in terms of a game $G_{x=y}$ played between two players: `Equal` and `Not-equal`. If both x_1 and y_1 are empty, then `Equal` wins; if one of them is empty and the other nonempty, then `Not-equal` wins. Otherwise, `Not-equal` picks a member x' of x or a member y' of y and `Equal` must reply with a member y' of y (if `Not-equal` had picked x' of x) or a member x' of x (if `Not-equal` had picked y' of y.). Then, they play $G_{x'=y'}$. It is routine to prove (by well-founded induction on $\in$) that $x = y$ iff player `Equal` has a winning strategy in $G_{x=y}$.

We will return to this game-theoretic imagery later.

So if we are to exploit properly the fact that our sets all live inside the cumulative hierarchy, then we have to be able to 'take sets apart'. This in turn means that we have to have an anti-wand. I like to think of the anti-wand as a kind of one-armed bandit. You put x into the slot and pull the handle, and after a brief clattering sound, the tray at knee level fills with members. The tray contents do not constitute a set, and they fail to constitute a set in exactly the same way that the lasso contents fail to constitute a set. The lasso contents fail to be a set because they have not been turned into one yet, and the tray contents fail to be a set because they are the result of taking the set apart. Notice also that the lasso contents and tray contents have no internal structure: in that respect, they are already a bit like sets. Notice, too, that the players never pick elements of *sets*; they pick elements from *presets*: lasso contents/tray contents.

Of course what is going on here is that we are thinking of sets as a datatype that is equipped with what the theoretical computer scientists call *constructors and destructors*. The one-armed bandit is simply the destructor function, and the wand is the corresponding constructor – the *sole* constructor as it happens. In contrast, the datatype of formulae has lots of constructors: $\wedge$, $\vee$, and so on, but the cumulative hierarchy has only one.

But notice that there is nothing in the idea of sets as conceived *iteratively* that says there should be only one constructor. Formulae are conceived iteratively and have more than one constructor: the fact that a datatype is conceived iteratively tells us nothing about how many constructors it has.

If there is nothing in the iterative conception of a datatype that tells us how many constructors that datatype is to have, perhaps we should consider the possibilities of other ways of constructing sets – iteratively. The umbrella this venture can shelter under is *Conway's principle* adumbrated in the appendix to part zero of Conway: "Objects may be created from earlier objects in any reasonably constructive fashion." (p. 66).

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

2. The two-constructor case. We can modify the cumulative hierarchy construction to one where – each time we create a new set by lassoing – we also create a companion to it which is to be its complement. As before, once you have lassoed a collection, you touch it with a wand and you obtain a set. However, this time you have a second wand, and as well as touching the lassoed collection with the first wand to get a set, you touch the lassoed collection with the second wand and you get the *complement* of the set made by the first wand. You ordain that the complement of x is to contain all the things so far created that are not in x, and additionally to contain all the sets as yet unborn.

Let us consider the first few stages of this construction to get clear in our minds the difference between this construction and the cumulative hierarchy construction.

Stage 0

At the outset, the world is empty, so when I throw out my lasso I catch nothing. Thus, the two wands give me $\emptyset$ and its complement V.

Stage 1

The world now contains two things: $\emptyset$ and V. So when I throw out my lasso, I might catch

(i) nothing, as before;
(ii) $\emptyset$, in which case the two wands give me $\{\emptyset\}$ and $V \setminus \{\emptyset\}$;
(iii) V, in which case the two wands give me $\{V\}$ and $V \setminus \{V\}$; or
(iv) both $\emptyset$ and V, in which case my two wands give me $\{\emptyset, V\}$ and $V \setminus \{\emptyset, V\}$.

Stage 2

So now the world contains 8 things: $\emptyset$, V, $\{\emptyset\}$, $V\setminus\{\emptyset\}$, $\{V\}$, $V\setminus\{V\}$, $\{\emptyset, V\}$ and $V\setminus\{\emptyset, V\}$.

Stage 3...

And now we have to say 'and so on ...' because (as the reader can calculate) at the next stage, we will have 512 things! Of course we iterate this through all the ordinals just as one iterates the cumulative hierarchy.

Sets created by the first wand are *low* sets, and sets created by the second wand are complements of low sets or *co-low* sets for short. Not hard to see that no low set can be self-membered, and every co-low set must be self-membered. There are a couple of important (because missable) trivialities to emphasise at this stage.

1. Our sets – of the new flavour as well as the old – will have birthdays or ranks, just as under the old dispensation.
2. *Low* is **not** the same as *well-founded*: every well-founded set is low, but there are plenty of low sets that are not well-founded: $\{V\}$ for one. However, it is true that no set created by the first wand can ever be a member of itself. Indeed, we can easily even establish that $\in$ restricted to first-wand sets is well-founded. (Consider the function that sends each set to the stage at which it is created.)

I shall discuss in the section 'Church–Oswald models' the question of what axioms for set theory are suggested by this construction. However, it may be worth noting at this stage that one can give for this construction the same kind of argument that it generates no paradoxical objects as one can give for the cumulative hierarchy construction. For example, consider the possible existence of the Russell class. If it ever got created, it would have to be either a low set or a co-low set. It cannot be a low set, since we keep on creating new sets that are not members of themselves: no low set is a member of itself after all. And the complement of the Russell class cannot be low, since we keep on creating sets that

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

are members of themselves: every co-low set is a member of itself. So the Russell class is neither low nor co-low, so it never gets created.

2.1. Set equality in the two-constructor case. Recall the game $G_{x=y}$ played by `Equal` and `Not-equal` in the cumulative hierarchy setting, and consider how it works in the new two-wand setting.

In the cumulative hierarchy setting, we never had to worry about infinite plays. The point is not that one player or the other always had a strategy to force a win after a finite play; the point is the stronger one that no infinite plays were possible at all. The situation here is very different.

Suppose we leave the termination conditions the same as in the cumulative hierarchy, as on page 3.

Consider the game $G_{V=V}$ in which the two players test whether or not $V = V$. The answer is 'yes', so `Not-equal` should lose the game. However, he can keep on postponing defeat by repeatedly playing V, a move to which `Equal` is forced to reply V every time.

(Obviously the same holds not just for V but for any self-membered set.) This is an infinite play that will happen even if both players know what they are doing, and it shows that at least some infinite plays have to be classed as wins for player `Equal`.

So do we want *all* infinite plays to be wins for player `Equal`? This might not be such a bad idea: it is simple to check that the relation "`Equal` has a strategy in the game $G_{x=y}$ that enables her to avoid defeat in finitely many moves" is a congruence relation for $\in$.[1] This – it could be argued (Forster, 1983) – is the only absolutely indispensable condition for a candidate for the role of 'equality between sets'. So it is not an obviously absurd idea to rule that all infinite plays are wins for player `Equal`. If we rule that all infinite plays are wins for player `Equal`, we are adopting what is known in the trade as the *axiom of strong extensionality*. This axiom is the formalisation of the view that two sets should be identical unless there is a good finite reason for them not to be.

However, this policy results in immediate disaster.[2] Consider the two sets $V \setminus \{V\}$ and V – both of them authentic products of the two-wand construction – and the game played to distinguish them. For his first move, player `Not-equal` clearly picks V from V – after all, it is his only hope – but then `Equal` picks $V \setminus \{V\}$ from $V \setminus \{V\}$, and the two players are back in exactly the situation they started in! Clearly `Equal` can postpone defeat for ever. Annoyingly we want *this* infinite play to be a win for player `Not-equal`!

Malitz's example appears to be telling us that if we have strong extensionality, then not both V and $V \setminus \{V\}$ can be sets: as we have just seen, `Equal` wins the game $G_{V=V\setminus\{V\}}$ by postponing defeat indefinitely.

For years, I thought this was something to do with the nonexistence of complements of singletons in the set theory GPC (called 'generalised positive comprehension'), which I am *not* proposing to discuss here (but see Esser, 1999, 2004, and references therein), but now I think this is mere coincidence.

I suggest instead that the significance of these two difficulties – V and Malitz's example – is that we have not made the necessary changes to $G_{x=y}$ in moving from the cumulative hierarchy case to the two-wand case. To accommodate the two-wand picture and still give a recursive account of set equality, we have to modify the rules of $G_{x=y}$ to take account of the fact that sets can now be constructed in more than one way. The rules for the two-constructor case should have been as follows.

---

[1] Strictly it is a congruence relation for the (infinitary) operation performed by the lasso.

[2] I am indebted to Isaac Malitz for drawing my attention to this example.

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

> If one of x and y is low and the other co-low, then `Not-equal` wins at once. If they are both low, `Not-equal` picks a member (x' or y') of one of them and `Equal` replies with a member (y' or x') of the other. If they are both co-low, `Not-equal` picks a member (x' or y') of one of $V \setminus x$ and $V \setminus y$, then `Equal` replies with a member (y' or x') or the other. They then play $G_{x'=y'}$.

The rationale for this is: since we now have *two* ways of constructing sets not one, then in order to tell when two sets are the same or distinct, we have to know what constructors the two sets arise from at top level and deconstruct them using the right destructor. We actually – all along – had to deconstruct a set for the $\in$-game, but since (in the cumulative hierarchy picture) there was only ever one constructor, we did not have to ask which one was being used. This time there are two, and we do have to ask. Readers should satisfy themselves that this process (the game) is terminating in exactly the same way and for the same reasons that the original game, $G_{x=y}$, was terminating in the cumulative hierarchy case: at each recursive call, we are considering objects with earlier birthdays. The situation is exactly like that with the construction of formulae. Two formulae are the same if they have the same list of immediate subformulae and have the same principal connective. We need to know the top-level connective of a formula when trying to parse it. Just as the truth-conditions of a formula depend on how it is constructed, so the identity conditions of sets depend on how *they* are constructed: a set is constructed from lasso contents *plus an operation.* This development can be easily generalised to encompass as many constructors as we like. Two sets are the same iff they are constructed using the same constructor on the same lasso contents. As long as no set can be made in more than one way,[3] we can continue to claim that identity between sets can be given the same recursive narrative as that given by the cumulative hierarchy and – who knows – even if it turns out that some sets are manufacturable in more than one way, we might yet find a way of coping.

So let us summarise. The claim is that one can painlessly generalise the construction of the cumulative hierarchy by postulating other constructors, and thereby building other sets, whose membership relation can be ill-founded.

However – despite their ill-foundedness – all these new sets nevertheless have well-founded structure of some kind, and it is in virtue of this recursive structure that we can continue to give a recursive account of set identity – enabling us to remain onside with the Quinean police[4] in exactly the way we were when we only had the cumulative hierarchy.

The key to making sense of this startling claim is to be found in teasing apart the two roles played by the membership relation $\in$. In the cumulative hierarchy picture, it is both (i) the membership relation between the sets we have created and (ii) the relation of *ontological priority* between those sets. (Strictly it is the ancestral or transitive closure of $\in$ that is the relation of ontological priority rather than $\in$ itself.) In Forster (2003), I call the ontological priority relation between objects in a recursive datatype the *engendering relation* of that datatype. It is not hard to show that the engendering relation of a recursive datatype is always well-founded and – being well-founded – is available to be exploited for transfinite algorithms for testing equality (and other relations).

---

[3] If we had a power set constructor as well as a finite-set constructor, then any power sets of a finite set would be constructed in more than one way: (i) by the power set constructor and (*ii*) – since it is also finite – by the finite-set constructor

[4] Who have "No entity without identity" tattooed on their knuckles so that it is the last thing you see before they . . .

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

We are now in a position to see that – although our sets have to form a recursive datatype, so we can have a recursive characterisation of set equality as in the two cases we have seen – the engendering relation of our recursive datatype of sets does not have to be $\in$ (or rather, its transitive closure). In the case of the cumulative hierarchy, the engendering relation happens to be $\in$, but in the two-wand case, it does not: it is the (rather cumbersome) relation which is the ancestral (transitive closure) of

"$(\text{low}(y) \wedge x \in y) \vee (\neg\text{low}(y) \wedge x \notin y)$"

. . . which is actually equivalent to

$$x \in y \longleftrightarrow \text{low}\,(y). \tag{3}$$

The role played by $\in$ in the cumulative hierarchy case is now played by the engendering relation of the new rectype: the relation (3) displayed above.

3. More wands. The two-wand case discussed above is of course only the simplest of a range of possible ways of complicating the construction that gave us the cumulative hierarchy – while remaining within the iterative conception of sets broadly understood. It is possible to spice up the construction so that – for example – every set not only has a complement but also generates a principal ultrafilter: $\{y : x \in y\}$. Also, if $\sim$ is an equivalence relation definable by quantifying only over low sets, one can add $\sim$-equivalence classes for all low sets (see Church, 1974). However, I will not be pursuing these possibilities since they do not serve to make any metaphysical points not already made by the two-wand case.

There are four attractive features of the cumulative hierarchy which the two-wand construction also has. They are the second-order categoricity of the corresponding set theory, the recursive decidability of equality, the availability of a natural notion of restricted quantifier, and the possibility of independence proofs by forcing.

3.1. Second-order categoricity. The two-wand picture has the same second-order categoricity as the one-wand picture. This, again, is just a result of the recursive nature of the definition. In second-order logic, one can prove that any recursive datatype whatever – even of infinite character – is unique and its second-order theory is categorical. The categoricity of second-order ZF is merely a special case of this – and the categoricity of the second-order version of the theory corresponding to the two-wand construction ('CUS' which we will see below) is another special case.

3.2. Equality. The equality relation between sets must be a congruence relation for $\in$, and strong extensionality is the idea that it should be the maximal congruence relation for $\in$. It could be argued that *prima facie* the only thing we know about equality between sets is that it is a congruence relation for $\in$. If there are lots of such congruence relations, then we have to decide which of them is equality. By assuming strong extensionality, we cut down the number of candidates to one and thereby simplify the decision mightily. Hence the attraction of an axiom of strong extensionality.

Strong extensionality is straightforwardly true in the cumulative hierarchy: given any two congruence relations for $\in$, we prove by induction on $\in$ that they agree. This is probably what makes strong extensionality attractive to us. But an analogous result holds in any iterative construction. It is just that in the more general case, the induction is not on $\in$ but on the engendering relation. In the two-constructor case, the axiom of strong extensionality is precisely the claim that equality is defined by the game $G_{x=y}$.

3.3. Restricted quantifiers. Any recursive datatype gives us a notion of restricted quantifier. In the arithmetic of $\mathbb{N}$, it is $(\forall x < y)$ and its dual. In the cumulative hierarchy, it is

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

$(\forall x \in y)$ and its dual. These restricted quantifiers are well-behaved because the relation that restricts the quantifier $- \in$ or $< -$ is the relation of *ontological priority*. There is always a well-defined notion of ontological priority with a recursive datatype: it is just the engendering relation. (There is no good notion of restricted quantifier in the reals or the rationals for example.) When we put it like this, it becomes clear what the appropriate notion of ontological priority is for the two-wand case: our restricted quantifiers will be $(\forall x\ R\ y)$ and $(\exists x\ R\ y)$ (where R is the relation (3) above).

One feature one expects restricted quantifiers to have is that one should be able to 'pull out' unrestricted quantifiers from within their scope and thereby prove a normal form theorem to the effect that every formula is equivalent to one that has no unrestricted quantifiers within the scope of a restricted quantifier. In the cumulative hierarchy picture, this is secured by means of the axiom scheme of collection. It is simple to check that the same normal form theorem holds for the new version of restricted quantifier that we have in the two-wand case. Suppose $(\forall x R y)(\exists z)(\phi)$. There are two cases to consider, y low and y co-low. If y is low, we use collection for low sets to obtain a low Z such that $(\exists Z)(\forall x R y)(\exists z R Z)(\phi)$. If y is co-low, we use low collection, this time on the complement of y.

This is all right as long as we have low collection. Does low collection follow from low replacement? Yes, because the usual argument using witnesses of minimal rank works.

It is not hard to see how one would extend the proof to more complex iterative constructions.

3.4. Forcing. Forcing is possible in the two-wand construction. The recursions on $\in$ that feature in the truth lemma and the definition of forcing for atomic formulae are replaced by recursions on the relation (3). Also $\check{a} =: \{\langle \check{b}, 1\rangle : b \in a\}$ if a is low. If a is co-low, then we must have $\check{a} =: V \setminus \{\langle \check{b}, 1\rangle : b \notin a\}$.

$$\check{a} =: \begin{cases} \{\langle \check{b}, 1\rangle : b \in a\} & \text{if } a \text{ is low} \\ V \setminus \{\langle \check{b}, 1\rangle : b \notin a\} & \text{if } a \text{ is co-low.} \end{cases}$$

Similarly,

$$a^F =: \begin{cases} \{b^F : (\exists p \in F)(\langle b, p\rangle \in a)\} & \text{if } a \text{ is low} \\ V \setminus (V \setminus a)^F & \text{if } a \text{ is co-low.} \end{cases}$$

Notice that there can be no co-low poset (the graph of a partial order on a co-low set would not be co-low), so all sets of conditions are low sets as before.

4. Objections. It is quite unrealistic to expect this novel picture to be welcomed with open arms; the cumulative hierarchy picture is far too well-entrenched. Boolos (1971) quotes Russell as saying that the other conceptions of set are not "such as even the cleverest logician would have thought of if he had not known of the contradictions."[5] Boolos has the

[5] One could equally wonder whether the iterative conception of set is the kind of thing even the cleverest logician would have thought of if he had not known of the contradictions. Since the discovery of the paradoxes predates not only NF and ZF but even all the articles giving expression to the cumulative hierarchy, we shall never know, but surely it is a safe bet that without the contradictions even the cleverest logicians would have gleefully forged ahead with naive set theory. One is reminded of Quine's (1969) remark on p. 242 of *Set theory and its logic* that "[impredicative definition] is hardly a procedure to look askance at, except as one is pressed by the paradoxes to look askance at something."

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

grace to admit "There does not seem to be any argument that is guaranteed to persuade someone who really does not see the peculiarity of a set's belonging to itself, or to one of its members etc., that these states of affairs are peculiar." Despite this, objections will surely be raised. Let us try to anticipate some of them.

4.1. Sets constituted by their members. Those of us with students have probably been telling them for years that what distinguishes sets from all other mathematical structures that have elements (lists, multisets, etc.) is that

Sets are uniquely characterised by their members. (2)

But now we find that the only sets that are defined by their members are sets created with the first wand. This seems to be telling us that the objects created by other wands – legitimate kosher extensional objects though they are – are not *sets* since they are not constituted by their members. But is this really the moral? I think not. Observation (2) is not so much a *definition* of set as a very useful *contrastive explanation*: a device for explaining sets to computer science students who have encountered other datatypes like lists and multisets. It captures the contrast between first-wand sets *and other first-wand objects* such as lists and multisets. It is the feature that one tries to point out once the first-wand nature of the entity has been agreed on in advance. The purpose of "Sets are constituted by their members" is not to make a contrast with "Some sets are constituted by their non-members" but with "Lists are constituted by their members *plus what you do to them*". (2) is really no more than an attention-getting way of saying that $\in$ is extensional. And extensionality of $\in$ does not rule out co-low sets.

4.2. The end of time. It will be objected that we do not know what the members of the complement objects are until the end of time. It is admitted that we know in some sense what the members *will* be, but this knowledge is of the set's *intension*, not its *extension* – and sets are supposed to be extensions!

The obvious riposte to this would appear to be the observation that we know what sets are members of what once the construction is halted at the end of time. We discover what extensions correspond to the intensions once we reach the end of time. How good a reply is this?

Not very. After all we never do reach the end of time. The process of creating new sets is everlasting, and not only will we be engaged on it as long as we have breath in our bodies, but when we die someone else will take over. We never reach a stage at which we down tools and start doing something else (such as some mathematics – which is partly what this was supposed to be for, let us not forget) because the recursion tells us at any stage to throw out the lasso and make more sets. Every stage is a stage at which you make more sets.

But notice that this argument can be used to equal effect against the cumulative hierarchy picture of sets too. The V_α's never stop either and for the same reason. What is to be done about this? What is the ZF-istes' reply?

They have at least two. One is "*Our* sets are all right from the moment they are born; a second-wand set isn't all right until the end of the construction." This seems to be an important difference.

Another reply is: since you cannot wait until the end of time (because there is no end of time), what you do is call a halt at some bounded stage.

But if you halt the two-wand construction analogously at some bounded stage, you *do* know what the members of all your sets are.

Perhaps that was not really what the ZF-iste meant: perhaps, rather than calling a halt to the process, simply interleave it with the execution of some mathematics. This seems

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

to have some bite: it makes the co-low sets look more intensional, in that one can be acquainted with them without knowing exactly what they contain: more might get put in as soon as you turn aside from doing the mathematics to constructing some more sets. But then, plenty of things like $\mathbb{R}$ and $\mathcal{P}(\mathbb{N})$ are intensional in parallel ways. The ZF-istes do not object to that, so presumably it is not a problem.

4.2.1. What is it an argument against?. The end-of-time argument swims so naturally and swiftly to the front of the minds of those used to the cumulative hierarchy that there is a tendency for them to overlook its other possible applications. Is this an argument against thinking of these objects as sets? That is certainly the use to which it is put. But surely – if it has any force at all – it has more force than that. It is not just an argument against the idea that these things are *sets*; it is an argument against the idea that they are legitimate mathematical objects *at all*. However, they surely are legitimate mathematical objects in virtue of Conway's principle. Careless or nervous readers might be spooked by the appearance of the complementation operator in the two-wand construction into thinking that it is not a monotone construction, but this is merely a trick of the light: the construction is perfectly monotone.

There may be room for debate about whether they are entities of the kind legitimated by Conway's principle.

4.3. How many wands?. If we are to admit that the cumulative hierarchy does not exhaust the collection of sets iteratively conceived, then how many constructors do we have to acknowledge? Once we have dropped our guard to the extent of allowing two wands, might there be no end to it and we would become bogged down in discussions of whether one constructor or another is legitimate. Might not it just be simpler never to let these people through the door in the first place?

My reply: it is an interesting question, but it is hardly a problem. It is the problem – if problem it be – confronting the man who tunnels into the cellars of a Château in Bordeaux and cannot work out which bottle to open first. I would like to have problems like that.

More seriously though, the thought that admitting other recursive constructions might give rise to difficulties later – and that a line will have to be drawn – is not *by itself* an argument for drawing the line so close to home that the two-constructor case is excluded. It may be a motivation to look for a significant difference between the one-constructor and the two-constructor cases, but it does not itself constitute such a difference, and in any case that search is – or soon will be – well under way anyway.

Here is one difference that it does not make: even if the two-wand story is not as metaphysically satisfying as the one-wand story, it is just as good from the point of view of arguing for consistency. The cumulative hierarchy motivates the adoption of a suite of axioms (ZF) for which it constitutes an obvious model. The two-wand construction provides a similarly convincing consistency proof for the set of axioms that it motivates. As we shall see later, one can give a consistency proof of this theory relative to ZF, but there is no need to, since the theory can be justified in the same way that ZF can. Accordingly objectors who wish to view the sets produced by the two-wand construction as spurious must think that there is more to the existence of mathematical objects than freedom from inconsistency.

5. Church–Oswald models. The two-wand construction (and its more complicated congenors) naturally gives rise to a body of axioms for set theory and in very much the same way that the cumulative hierarchy gave rise to the axioms of ZF, and in this section, we will see what axioms are suggested by the two-wand construction.

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

Interestingly, these theories also admit fairly direct consistency proofs relative to ZF by a method due independently and simultaneously to Church and Oswald. A thorough technical treatment of these methods is to be found in Forster (2001), but we will sketch one example – the two-wand construction – briefly here.

Working with a model $\mathfrak{M} = \langle V, \in \rangle$ of ZF(C), fix once for all a definable bijection k between V and $V \times \langle 0, 1 \rangle$, which must satisfy the property that the ranks of $k(\langle x, 0 \rangle)$ and $k(\langle x, 1 \rangle)$ must both be greater than the rank of x. Now define a new model with the same domain and a new membership relation by

$$x \in_{new} y \longleftrightarrow \begin{cases} k(y) = \langle y', 0 \rangle \text{ and } x \in y', \text{ or} \\ k(y) = \langle y', 1 \rangle \text{ and } x \notin y'. \end{cases}$$

It is now a laborious but relatively straightforward technical exercise to check that the new structure we have just characterised is the same up to isomorphism as the result of the two-wand construction described earlier. Even those with little appetite for this kind of detail will find it easy to check that if the second component of $k(x)$ is 0, then x is low and if it is 1, then x is co-low.

Now just what have we created a model *of* by these two processes?

The well-founded sets created by the two-wand construction at stage α are clearly the same as the well-founded sets we created in the cumulative hierarchy by stage α, so according to this construction, the well-founded sets are a model of ZF (if that is what it was we concluded the cumulative hierarchy is a model of).

Clearly we have separation for low sets, just as we had for all sets in the cumulative hierarchy case.

Whatever reason caused us to believe that the axiom scheme of replacement holds in the cumulative hierarchy will cause us to believe in this context that a surjective image of a low set is a low set. This tells us – for example – that the axiom of pairing holds: whatever x and y are, $\{x, y\}$ is going to be a low set whose birthday is the first day after the birthdays of x and y. Every low class will be a set. This of course also gives us an axiom scheme of replacement for low sets: a surjective image of a low set in a function class is clearly a set – indeed, it is a *low* set – for precisely the reason that the cumulative hierarchy is a model of replacement: the argument is the same.

Binary unions – $x \cup y$ – will always exist: if x and y are both low, then $x \cup y$ is low too; if x and y are both co-low, being the complements of the low sets x' and y', respectively, then $x \cup y$ is the complement of the low set $x' \cap y'$. If one is low and the other not (as in the case where x is the complement of the low set x') and y is low, then $x \cup y$ is the complement of $x' \setminus y$, which is a low set by separation.

The power set of a low set will be a low set. The power set of a co-low set will be neither low nor co-low, and so the axiom of power set holds only for low sets in the models obtained by this method. Sumset is different. Notice that every co-low set must contain a co-low member: every set containing only low sets must be of bounded rank. Any superset of a co-low set is co-low, so every sumset of a co-low set is co-low. A sumset of a low set is low unless it has a co-low member.

The result is that this construction gives us a model of a theory whose axioms are

1. Extensionality;
2. Existence of $x \cup y$;
3. Existence of $\{x, y\}$;
4. Existence of $V \setminus x$;

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

5. The (externally) well-founded sets form a model of ZF;
6. Sumset;
7. Replacement for low sets.

The fifth and seventh axioms are of course schemes. The first four axioms are a system called NF_2. The fifth axiom is a reflection of the fact that (modulo some small print) the well-founded sets of the two-wand construction are precisely the sets in the cumulative hierarchy. This is roughly the theory CUS of Church (1974).

So there are two ways of obtaining models for theories like those discussed in Forster (2001) and Church (1974). One can (i) elaborate them directly by iterative constructions that parallel the construction of the cumulative hierarchy and (ii) obtain them by coding up novel membership relations within a model of ZF. As far as we can tell, (ii) preceded (i) but there appears to be no significance to this fact. Did (ii) really precede (i)? Admittedly it would be very odd for Church to have thought as hard as he did about CUS without discovering the two-wand construction. But would it not have been odder still for him to have discovered the two-wand construction and then say nothing about it, instead proceeding directly to the Church–Oswald model construction? Perhaps not: it may be that Church's formalist leanings predisposed him to see this material entirely through syntax and that he did not know the two-wand construction. Of course it may also be that he knew it but – again, because of his formalist leanings – felt that wands were not the right instruments for presenting this idea.[6]

5.1. The significance of the Church–Oswald interpretation. So the two-wand iterative constructions can all be coded up inside the cumulative hierarchy by the techniques of Church and Oswald. It is a good question what the metaphysical significance of this coding might be. Does the existence of the coding confirm our belief that

> These more general iterative constructions are authentic constructions of genuine mathematical objects *Horn* 1

Or does it on the other hand rather explain why

> These more general iterative constructions are merely pointless epiphenomena *Horn* 2

The dilemma sporting these two horns is of course Moore's *paradox of analysis*.

Those who want to take the second horn have to have a reason for believing in the legitimacy of the cumulative hierarchy construction that goes beyond the legitimacy that the cumulative hierarchy acquires from being a sensible recursive datatype declaration. So far – as we have seen – the iterative conception and the cumulative hierarchy have always been conflated, so arguments for the first have been mistaken for arguments for the second. Hitherto this has always had the result that arguments of the kind now required by the second-horn votaries have never been specifically identified. Now at least, votaries of the second horn do at any rate know what it is they have to establish.

5.2. Forti–Honsell antifoundation. It is natural in this discussion to wonder whether or not there is a connection with Forti–Honsell style antifoundation axioms (Forti & Honsell, 1983). It is possible to think of Forti–Honsell universes as recursively constructed: at stage α, add all subsets of what you have so far and then – before proceeding to stage $\alpha + 1$ – add all sets whose set pictures you have just thereby added. This observation is

[6] Thanks to the referee and Isaac Malitz and Flash Sheridan for discussions on this point.

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

behind the proof that in universes of this kind, every set is of the same size as a well-founded set. We would have to have a constructor that created sets from set pictures. On the face of it, no such constructor would be injective and we would have the problem mentioned earlier of 'sets being created in more than one way'. There will surely be ways round this problem (in this case), but we have no space to cover them here.

6. Envoi: Why considering the two-wand construction might be helpful. Naturally, I prefer the first horn of the dilemma on page 108. The reason why we should extend a warm welcome to theories like CUS is that they strive very hard to do something ZF does not even consider attempting. They represent an endeavour to reason seriously and logically about the genuinely infinite, The Unbounded. Attempts to grasp the nettle and take The Unbounded seriously (one thinks of Cantor and Dedekind) have in the past been attended with success and have engendered some useful new mathematics. It might be the same again.

But do they in fact represent such an endeavour? Against this can be made the point that since the two-wand construction can always be coded inside the cumulative hierarchy, it can have no new mathematics to tell us. It is not clear how strong a point this is: not everybody is of the view that the interpretability of Forti–Honsell antifoundation (AFA) into ZF+ foundation means that AFA contains no new mathematics. In any case, even if one does conclude that the interpretability of the two-wand construction inside the cumulative hierarchy means that the the two-wand construction contains no new mathematics (on the grounds that everything it has to say can be encoded as facts about well-founded sets), that does at least mean that some of the results about well-founded sets are encoded assertions about big sets and should therefore be seen in that light: insights into large sets were buried inside the cumulative hierarchy all along. The fact that every hereditarily finite set is well-founded is a way of encoding the fact that every set of cofinite sets is included in one of its members. One is reminded of the tale of the two dreamers from the *One Thousand and One Nights* (night 351).[7]

7. Acknowledgments. Thanks are due to my Auckland students and to Flash Sheridan, Fred Kroon, Isaac Malitz, and Herb Enderton.

BIBLIOGRAPHY

Barwise, J., & Moss, L. (1996). *Vicious Circles*. Stanford, CA:CSLI publications.

Boolos, G. (1971). The iterative conception of set. *The Journal of Philosophy*, **68**, 215–231.

Church, A. (1974). Set theory with a universal set. In Henkin, L., editor, *Proceedings of the Tarski Symposium. Proceedings of Symposia in Pure Mathematics XXV*. Providence, RI: AMS, pp. 297–308. Also in *International Logic Review*, **15**, 11–23.

[7] A man in Cairo lives in a house with a fig tree in the garden. He has a dream in which he is told that a great fortune awaits him in Isfahan, whither he travels at once. There he gets into various scrapes, and when he is arrested the chief of police asks him what he is doing in Isfahan. On being told the story the Chief of Police says "Fool! I have this recurrent dream in which I am told of a house in Cairo [and he proceeds to give an exact description] where there is a fortune in gold coins buried under a fig tree. I pay no attention to this nonsense; you, on the other hand, travel half-way across the civilised world merely to make a spectacle of yourself. Begone." The Cairene goes back to his house and digs up the money.
The story is retold in Borges: *A Universal History of Infamy*.

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press

Conway, J. H. (2001). *On Numbers and Games*, 2nd edition. A. K. Peters, Natick, MA: Academic Press.

Esser, O. (1999). On the consistency of a positive theory. *Mathematical Logic Quarterly*, **45**, 105–116.

Esser, O. (2004). *Une théorie positive des ensembles. Cahiers du Centre de Logique*, vol. 13. Louvain-la-Neuve, Belgium: Academia-Bruylant. ISBN 2-8729-687-6.

Forster, T. E. (1983). Axiomatising set theory with a universal set. In Crabbé, M., ed. *"La Theorie des ensembles de Quine" Cahiers du Centre de Logique*, vol. 4., pp. 61–76. Louvain-La-Neuve, Belgium: CABAY. Available online from: `www.dpmms.cam.ac.uk/~tf/old_nf.ps`.

Forster, T. E. (2001). Church-Oswald models for set theory. In *Logic, Meaning and Computation: Essays in Memory of Alonzo Church*, Anderson, C. A. A. and Zeleny, M., editors, *Studies in Epistemology, Logic, and Philosophy of Science*. vol. 305. Dordrecht, Boston, and London: Kluwer.

Forster, T. E. (2003). *Logic, Induction and Sets, LMS Undergraduate Texts in Mathematics*, vol. 56. Cambridge University Press.

Forti, M., & Honsell, F. (1983). Set theory with free construction principles. *Annali della Scuola Normale Superiore di Pisa, Scienze Fisiche e Matematiche*, **10**, 493–522.

Quine, W. V. (1968). Speaking of objects. In *Ontological Relativity and Other Essays*.

Quine, W. V. (1969). *Set Theory and Logic*, 2nd edition. Cambridge, MA: Belknap Press Harvard.

CAMBRIDGE UNIVERSITY
CENTRE FOR MATHEMATICAL SCIENCES
CAMBRIDGE CB3 OWB, UK
E-mail: T.Forster@dpmms.cam.ac.uk

https://doi.org/10.1017/S1755020308080064 Published online by Cambridge University Press