The Liar

An Essay on Truth and Circularity

Jon Barwise and John Etchemendy
Stanford University

Oxford University Press
New York Oxford

Contents

3

The Universe of Hypersets

Set theory from Z to A

In both of our analyses of the paradox we take seriously the intuition that the propositions involved are genuinely circular. Since we are going to model propositions, situations, and facts with set-theoretic objects, it is extremely inconvenient to adopt a set-theoretic framework that precludes circular or nonwellfounded objects. The source of this inconvenience is simple. By far the most natural way to model a proposition about a given object is to use some set-theoretic construct containing that object (or its representative) as a constituent, that is, where the object appears in the construct's hereditary membership relation. But if we carry out this straightforward approach in a set theory based on Zermelo's cumulative hierarchy, we find ourselves inadvertently excluding the possibility of circular propositions. For the model of a proposition about another *proposition* will have to contain the latter's representative as a constituent, and the model of a circular proposition, one directly or indirectly about itself, will have to contain itself as a constituent. But the axiom of regularity, or foundation, bans sets that are members of themselves, or pairs of sets that are members of each other, and so forth, and so would block us from using such natural techniques of modeling propositions.[1]

[1]The axiom of foundation asserts that the membership relation is wellfounded, that is, that any nonempty collection Y of sets has a member $y \in Y$ which is disjoint from Y. This follows from the iterative conception by chosing any $y \in Y$ of "least rank," that is, a y that occurs as early in

There are various ways we could sidestep this problem within standard set theory, but they would involve us in complexities of considerable magnitude, ones completely irrelevant to the task at hand. Of course if worse came to worse, we could just give up set theory entirely as our working theory. If there had been no coherent alternative to the Zermelo conception of sets, one that admits circularity, we would probably have done just that. But Peter Aczel has recently developed an appealing alternative conception of sets, and with it a consistent axiomatic theory tailor-made for our purposes. Aczel's theory, based on an extremely natural extension of the Zermelo conception, is quite easy to learn and, once learned, lets us bring to bear all of the familiar set-theoretic techniques to the problem of modeling circular phenomena. We devote this chapter to an exposition of Aczel's theory, one that will allow the reader to follow the details of the rest of the book, as well as apply the theory in other domains.

To appreciate the intuitive appeal of Aczel's conception, let's first rehearse a common way of picturing ordinary sets. Consider, for example, the set $c_0 = \{a_0, b_0\}$ where $a_0 = \{\text{Claire, Max}\}$ and $b_0 = \{a_0, \text{ Max}\}$. There are many ways to picture this set, but one natural and unambiguous way is with the labeled graph shown in Figure 1. In this graph each nonterminal node represents a nonempty set, the set containing the objects represented by the nodes below it. For example the top node in the graph represents the set c_0, a set whose only members are the sets a_0 and b_0, and these latter sets are in turn represented by the nodes immediately below the top node. Note that the set represented by a node need not itself *be* the node, and indeed in Figure 1 we find two different nodes that each depict, and so are labeled by, the single set a_0. The bottom nodes in this example represent Max and Claire, neither of whom have elements, and so there are no nodes below them. The idea of such a graph, of course, is that the arrows represent the converse membership relation: an arrow from node x to node y indicates that the set (or atom) represented by y is a member of the set represented by x.

Notice that one and the same set may well be depicted by many different graphs. Consider, for example, the graphs in Figure 2.

---

the cumulative hierarchy as any other member of Y. This rules out circularity. For example, note that if $a \in a$ then the set $Y = \{a\}$ violates this assumption.

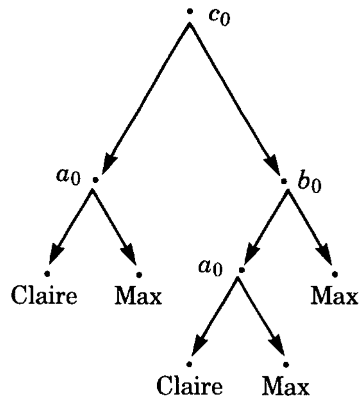

Figure 1

These are clearly different graphs since they have different graph-theoretic properties. For example, the three graphs have different numbers of nodes, and the first is a tree while the others are not. Still, as we've indicated by the labeling, they all depict the same set, the von Neumann ordinal three. Similarly, Figure 3 gives a different, and more economical depiction of our original set c_0. The differences among these graphs, for our purposes, amount to little more than the relative economy of nodes: Figure 3 has four fewer nodes than Figure 1, but it gives us a picture of exactly the same set.

In the same way, any set can be depicted by a graph. One canonical way to build a graph is to start with the desired set a and consider all of its "hereditary" members (members, members of members, members of members of members, and so on) as

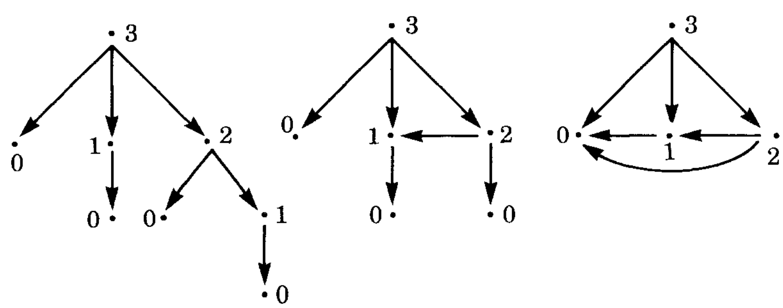

Figure 2

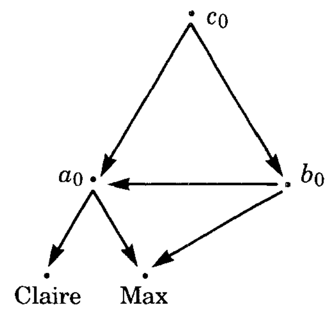

Figure 3

nodes of a graph. Then draw an edge from any set to each of its members. The resulting graph will depict the given set a (and in this case, the nodes actually *are* the sets depicted). This construction allows us to build a canonical graph for any set whatsoever. (And note that this does not presuppose that we are dealing with sets under the cumulative conception. Under any conception, sets give rise to graphs in this way.)

Exercise 7 Draw two graphs that represent the von Neumann ordinal four. Make the first graph a tree (on the model of the first graph in Figure 2), and make the second as economical as possible (on the model of the third graph in Figure 2).

Aczel's conception of a set arises directly out of the intuition that a set is a collection of things whose (hereditary) membership relation can be depicted, unambiguously, by graphs of this sort. The liberating element is that we allow arbitrary graphs, including graphs that contain proper cycles. Of course graphs with cycles cannot depict sets in the wellfounded universe.[2] Thus, for example, in Aczel's universe there is a set $\Omega = \{\Omega\}$, simply because we can picture the membership relation on Ω by means of the graph G_Ω shown in Figure 4. Furthermore, on Aczel's conception this graph *unambiguously* depicts a set; that is, there is only one set with G_Ω as its graph. Consequently, there is only one set in Aczel's universe equal to its own singleton.

[2]To see this, suppose we have a graph with a proper cycle. Take Y to be the set containing all the sets depicted by nodes that occur in the cycle. It is clear that no member of Y is disjoint from Y, thus violating the axiom of foundation.

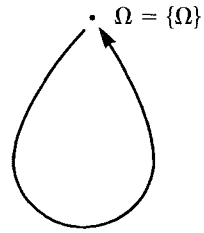

Figure 4

To take another example, let's consider the set $c = \{a, b\}$ where $a = \{\text{Claire, Max}, b\}$ and $b = \{\text{Max}, a\}$. The only difference between this set and our set c_0 above, is that b is an element of a, while b_0 is not an element of a_0 (and could not be, according to the cumulative conception). To get a graph of c, we can simply modify a graph of c_0, say the one given in Figure 3. Here, we need only add an edge from the node that represents a_0 to the node that represents b_0. The result, in Figure 5, is a graph of c.

The sets we get on Aczel's conception include all those in the traditional, wellfounded universe. But in addition to these, we get a rich class of nonwellfounded sets, or as we will sometimes call them, *hyper*sets. As we'll see, these sets behave in many respects just like the ordinary, wellfounded variety. But they permit the use of straightforward modeling techniques even when the phenomena modeled involve circularity.

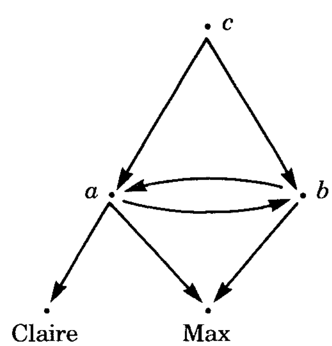

Figure 5

AFA

Let's flesh out Aczel's idea in more detail and describe the axiomatic theory explicitly. Actually, we will describe a variant of Aczel's theory that allows a collection $\mathcal{A}$ of "atoms" like Max and Claire, instead of just pure sets. This theory, then, will consist of all the usual axioms of ZFC set theory (modified in the usual way to admit atoms), except that the axiom of regularity is replaced by a strong form of its negation, called AFA, for (Aczel's) Anti-Foundation Axiom (with Atoms).

We said that on Aczel's conception a set is any collection of objects whose hereditary membership relation can be pictured by a graph. More precisely, a *graph* G is a set of nodes and directed edges, as usual. (Any set X can be a set of nodes, and any set $R \subseteq X \times X$ of ordered pairs from X can be used to represent the directed edges of a graph G. It is customary to write $x \rightarrow y$ to indicate that the graph contains an edge pointing from node x to node y.) If there is an edge $x \rightarrow y$ from node x to node y, then y is said to be a *child* of x. A node with no arrow starting from it is said to be *childless*. So, for example, in Figure 3 there are two childless nodes and three "parent" nodes, that is, nodes with children. In Figure 4, on the other hand, there is only one node, and it is a child of itself.

A *tagged graph* is a graph in which each childless node x has been "tagged" by an object $tag(x)$, which is either an atom or the empty set. Think of tagging as the process of simply writing the name of an atom or the empty set next to each childless node to indicate what it represents. More formally, a tagged graph is a graph G together with a function tag mapping the childless nodes of G into $\mathcal{A} \cup \{\emptyset\}$. (Note that if G has no childless nodes, as with Figure 4, then the totally undefined function suffices to tag the graph.) Aczel's basic idea is that once we have a tagged graph, we can use the nodes and edges of the graph to picture sets and set membership. To make this notion precise, we bring in the concept of a decoration for a tagged graph.

A *decoration* for a tagged graph is a function $\mathcal{D}$ defined on the nodes of the graph such that for each node x, if x has no children, then $\mathcal{D}(x) = tag(x)$, whereas if x has children, then

$$\mathcal{D}(x) = \{\mathcal{D}(y) \mid y \; is \; a \; child \; of \; x\}.$$

Each node x of G that has children is said to *picture* the set $\mathcal{D}(x)$. Thus we can think of the process of decorating a graph as simply continuing the process started by tagging the graph: we write next to each parent node a name of the set it depicts. We have in fact decorated all of the above graphs in just this way.

AFA can now be stated quite simply: it asserts that *every tagged graph has a unique decoration.* It is clear that this axiom conflicts with the Zermelo conception since, for example, the graphs in Figures 4 and 5 cannot be decorated with sets from the cumulative hierarchy. For example to decorate the graph in Figure 4, $\mathcal{D}$ would have to assign to the single node some set that contains itself. But there is no such set among the wellfounded sets.

There are two parts to AFA, existence and uniqueness. That is, part of what AFA asserts is that every tagged graph has a decoration. This guarantees the existence of all the sets we consider. However, equally important in applications is the uniqueness half, the assertion that no graph has more than one decoration. It is this part of the axiom that gives us a useful handle on the identity of nonwellfounded sets.

Consider, for example, the sets $a = \{\text{Max}, a\}$ and $b = \{\text{Max}, b\}$. Does $a = b$? The usual axiom of extensionality is useless in answering this question, for it asserts only that $a = b$ if a and b have the same members, which boils down to the assertion that $a = b$ if $a = b$. However, on Aczel's conception, it turns out that a is indeed equal to b since they are depicted by exactly the same graphs. To see this in detail, suppose we have a tagged graph G and a decoration $\mathcal{D}$ that assigns a to a node x of G, $\mathcal{D}(x) = a$. Consider the decoration $\mathcal{D}'$ just like $\mathcal{D}$ except that $\mathcal{D}'(x) = b$. A second's thought shows that $\mathcal{D}'$ must also be a decoration for G. But by the uniqueness part of AFA, we must have $\mathcal{D} = \mathcal{D}'$ and so $\mathcal{D}(x) = \mathcal{D}'(x)$; i.e., $a = b$.

On Aczel's conception, then, for two sets to be distinct there must be a genuine structural difference between them, one that prevents them from being depicted by the same tagged graph. This will be important in what follows, since the identity conditions on sets give rise to identity conditions on the various set-theoretic models that we construct below.

It's fairly easy to see which graphs can be decorated with wellfounded sets. Say that a graph G is *wellfounded* if for each nonempty subset Y of the nodes of G, some node in Y has no

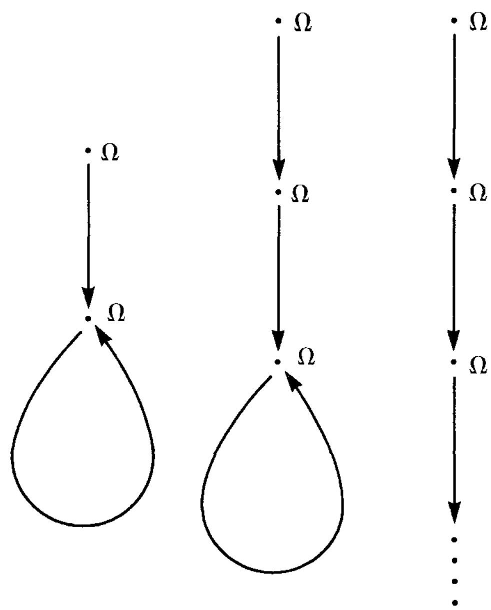

Figure 6

child in Y. Only wellfounded graphs can have decorations in the Zermelo universe; indeed the claim that no nonwellfounded graph can be decorated is just a reformulation of the axiom of foundation. Further, Mostowski's Collapsing Lemma[3] tells us that *every* wellfounded tagged graph has a unique decoration in the universe of wellfounded sets. So we can think of Aczel's axiom as extending this natural relationship between graphs and sets beyond the wellfounded.

It should once again be noted that under either conception, a set can in general be depicted by many different graphs. Figure 2 presented three different graphs of a wellfounded set, the von Neumann ordinal three. Similarly, Figure 6 gives a few additional graphs of the nonwellfounded set Ω. In this case, to see that each of these graphs depicts Ω, we need only note that all the nodes *can* be decorated with Ω, and so by AFA *must* be, the decoration being unique.

---
[3]See, for example, Kunen (1980), 105.

While Aczel's is quite a different conception from Zermelo's, it turns out that all the usual axioms of ZFC are true under this conception, except, of course, the axiom of foundation. This means that we can use all the familiar set-theoretic operations (intersection, union, power set, ordered pairs, and so forth) without any change whatsoever. Only when the axiom of foundation enters (as with inductive definitions, which we discuss in the last section of this chapter) do we need to rethink things.

Let's look at one more example, this time a bit more relevant for our purposes. Consider the English sentence

(ϵ) This proposition is not expressible in English using ten words

and the various propositions it can express. Let us use an atom E to represent the property that holds of a proposition just in case it is expressible in English using ten words. Suppose we were to model the proposition that p has the property E with the triple $\langle E, p, 1\rangle$, and the proposition that p does not have E with the triple $\langle E, p, 0\rangle$.[4] Recall that in set theory triples $\langle x, y, z\rangle$ are taken to be pairs of pairs $\langle x, \langle y, z\rangle\rangle$, that an ordered pair $\langle y, z\rangle$ is construed as the set $\{\{y\}, \{y, z\}\}$, and that 0 is represented by the empty set. Then we can see that a graph of our model of the proposition that p does not have E and a graph G_p of p are related as in Figure 7.

Suppose we want to represent the (intuitively false) circular proposition expressed by (ϵ) when "this proposition" is given the reflexive reading. This will be the proposition q that claims, of itself, that E does not hold. That is, we want $q = \langle E, q, 0\rangle$. By what we have just said, it suffices to take the special case of Figure 7 where the graph G_p is the whole graph. This is shown in Figure 8. Thus, the proposition we are after is modeled by the set assigned to the top node in Figure 8. There is exactly one such set in the universe of hypersets.

Exercise 8 Label the unlabeled nodes of the graphs in Figures 7 and 8.

Exercise 9 Show that Ω is depicted by all the graphs in Figure 9.

---

[4]To keep the graph simple, we are suppressing the atom *Prop* introduced in Chapter 2, page 28.

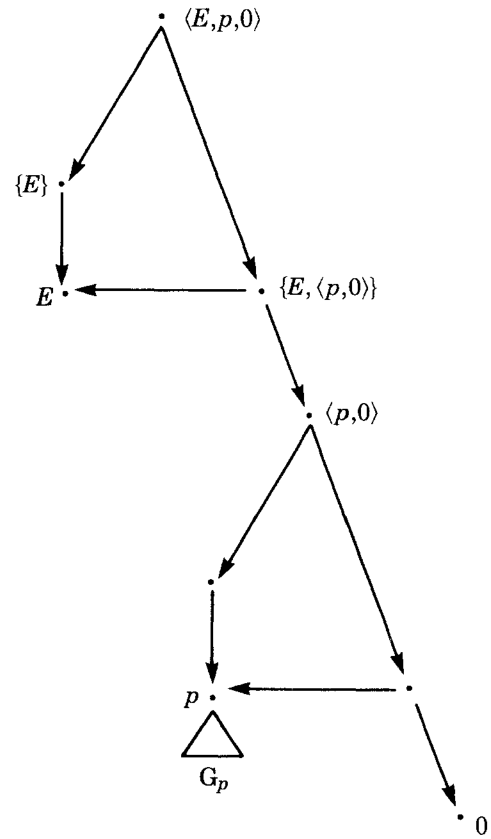

Figure 7

Exercise 10 Using AFA, show that there is a unique set a satisfying the equation

$$a = \{a, \emptyset\}.$$

Show that $a \neq \Omega$.

Exercise 11 Show that the graph shown in Figure 5 is nonwellfounded. That is, find a nonempty set Y of nodes of the graph such that every member of Y has a child in Y.

Exercise 12 Say that a graph is transitive if for each pair of edges $x \to y$ and $y \to z$ there is an edge $x \to z$. Similarly, say that a set a is transitive if $c \in b \in a$ implies $c \in a$. Show that a set is transitive if (but not only if) it is depicted by a transitive graph.

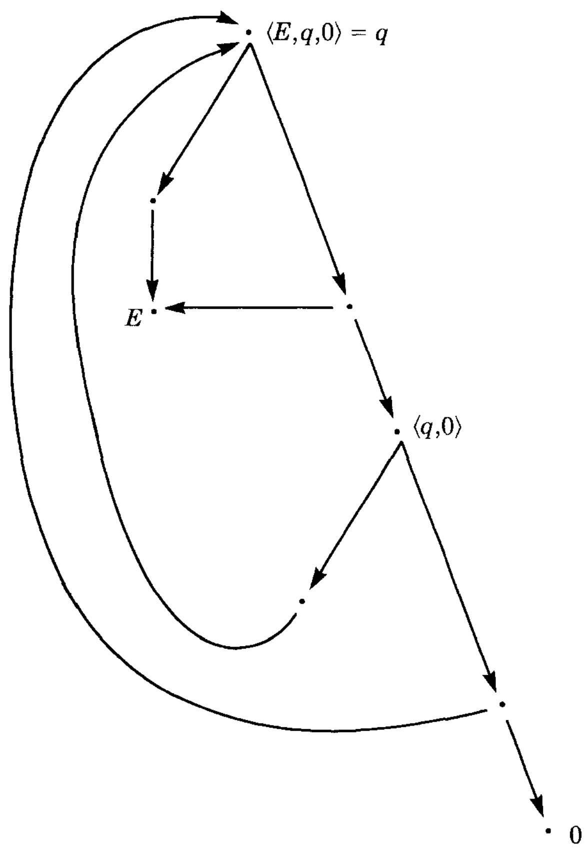

Figure 8

The transitive closure of a set a is the smallest transitive set containing a as a subset. Suppose that a node x of a graph pictures the set a. Show that the transitive closure of a is the set of all decorations of nodes appearing "below" x. (By "y is below x" we here mean that there is a path of arrows from x to y.)

The consistency of ZFC/AFA

There were really two sorts of set-theoretic paradoxes that threatened early, intuitive set theory: paradoxes of size and paradoxes like those engendered by the Russell set, the set z of all sets that

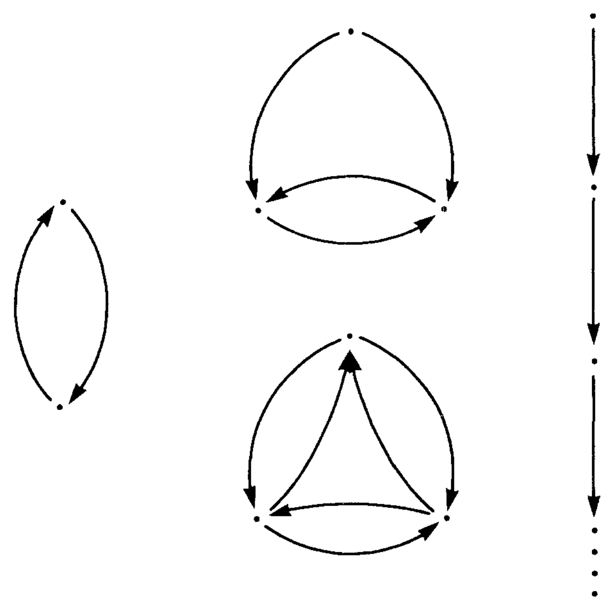

Figure 9

are not members of themselves. The Zermelo conception killed two birds with one stone. On the one hand, it gave us a way to conceptualize classes that are never collected into sets, and, on the other, it ruled out sets that are members of themselves. But as a reaction to the paradoxes, this latter move was really unnecessary. On Zermelo's conception the Russell "set" is actually the universe of all sets. And since this is a proper class, not a set at all, the familiar reasoning that derives a contradiction from the definition of z is blocked. But the set/class distinction is the key here, not the banning of self-membership.

On Aczel's conception, we still have the set/class distinction, only now there is a proper class of sets that *do* contain themselves, as well as a proper class that *do not*. (See Exercise 14.) In both cases there is no Russell set, only a Russell class. To obtain sets using the Russellian definition, the comprehension schema does not allow the earlier definition of z:

$$z = \{x \mid x \notin x\}$$

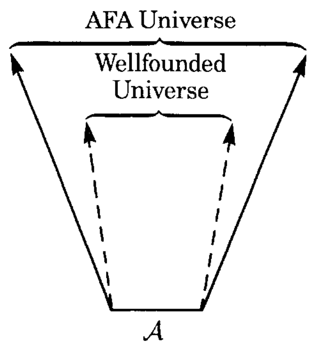

Figure 10

but rather requires that we introduce a parametric version of the definition:

$$z_a = \{x \in a \mid x \notin x\}.$$

What the Russell argument now shows is just that z_a can never be in the set a, whether or not a is wellfounded. The set z_a is said to "diagonalize out" of the set a.

Since we are working in the realm of the paradoxes, both set-theoretic and semantic, it is obviously important to be sure that our metatheory, ZFC/AFA, is consistent. Aczel has shown that it is.[5] Indeed, he has shown more. Working in ZFC$^-$ (ZFC without the axiom of foundation), Aczel shows how to canonically embed the universe of wellfounded sets into a universe satisfying ZFC/AFA, what we have been calling the universe of hypersets. We call this result the Embedding Theorem. Since the construction yields a model of ZFC/AFA, it shows that the theory is consistent, assuming of course that ZFC is. But it also shows that we can think of the universe of hypersets as a mathematical enrichment of the universe of wellfounded sets. Thus we can depict the relationship between the two as in Figure 10.

The situation here is entirely analogous to any number of similar cases in mathematics. For example, consider the relation between the real numbers and the complex numbers. The familiar model of the complex numbers as equivalence classes of pairs of reals yields a consistency proof of the theory of complex numbers

[5]See Aczel (1987).

relative to the theory of real numbers. But it also does something more: it shows us that the complex numbers can be thought of as an expansion of the reals.

The proof of the Embedding Theorem, while a bit tricky in detail, is simple enough to describe. First Aczel isolates an equivalence relation $\equiv_A$ on graphs which holds between two graphs just in case they represent the same set. For example, all the graphs in Figure 2 are $\equiv_A$, as are the four graphs from Figures 4 and 6. This allows each set in Aczel's universe to be represented by an equivalence class of graphs from the wellfounded universe. There is a slight hitch, though, since each set is actually depicted by a proper class of graphs, and to carry out the proof in ZFC$^-$ one has to work with sets. To do this Aczel borrows a trick of Dana Scott's, and represents each set b by the *set* G_b of those graphs of minimal rank in the cumulative hierarchy that depict it. Since every graph is, by the axiom of choice, isomorphic to a graph on some set of ordinals, G_b will always be nonempty.[6] Then, using the class of sets of the form $G_b/\equiv_A$, Aczel is able to show (1) that all the axioms of ZFC/AFA are true (using the natural interpretation of membership), and (2) that every wellfounded set is uniquely represented in the resulting model.

Aczel's proof shows that there is a sense in which AFA does not give rise to any new mathematical structures. One could always replace talk of the nonwellfounded sets in the AFA universe with talk of the structures $G_b/\equiv_A$, just as one could replace talk of complex numbers with talk of equivalence classes of pairs of real numbers, or replace talk of real numbers with talk of equivalence classes of Cauchy sequences of rationals. You could do any of these in principle, but it would be completely impractical, and ultimately misguided. As mathematical objects, the complexes are as legitimate as the reals, and the AFA universe is as legitimate as the universe of wellfounded sets. The fact that we can model one with the other does not make the latter more basic or more legitimate than the former.

Exercise 13 Recall the definition of the parametric Russell set z_a given above. What is z_Ω? Let c be the nonwellfounded set

[6] Notice that this observation also shows that we get the same AFA universe whether our graphs are drawn from the wellfounded universe or from the full AFA universe.

depicted in Figure 5, page 38. What is z_c? Let a be the set defined in Exercise 10, page 43. What is z_a?

Exercise 14 Show that for any set a, there is a set $b = \{a, b\}$. Show that distinct sets a thereby give rise to distinct sets b. Conclude that there is a proper class of sets which are members of themselves.

Solving equations

In addition to standard set-theoretic facts from ZFC, there is one simple consequence of AFA that we will use over and over in what follows, a result that allows us to assert that various sets exist without first depicting them with graphs.

Consider an "indeterminate" $\mathbf{x}$ and the equation

$$\mathbf{x} = \{\mathbf{x}\}.$$

This equation has a solution[7] in the universe of hypersets, namely Ω. Furthermore, since any solution to this equation would be depicted by the graph G_Ω, this equation has a unique solution in the universe.

Similarly, consider the following three equations in the indeterminates $\mathbf{x}$, $\mathbf{y}$, and $\mathbf{z}$.

$$\begin{aligned} \mathbf{x} &= \{\text{Claire}, \text{Max}, \mathbf{y}\} \\ \mathbf{y} &= \{\text{Max}, \mathbf{x}\} \\ \mathbf{z} &= \{\mathbf{x}, \mathbf{y}\} \end{aligned}$$

AFA tells us that these equations have a unique solution in the hyperuniverse, the sets $\mathbf{x} = a, \mathbf{y} = b$, and $\mathbf{z} = c$ pictured in Figure 5, page 38.

Aczel has a general result which allows us to find, for any system of equations in indeterminates $\mathbf{x}, \mathbf{y}, \mathbf{z}, \ldots$, say,

$$\begin{aligned} \mathbf{x} &= a(\mathbf{x}, \mathbf{y}, \ldots) \\ \mathbf{y} &= b(\mathbf{x}, \mathbf{y}, \ldots) \\ &\vdots \end{aligned}$$

[7] We use the term "solution" in exactly the same way as it's used in algebra. Below we will represent a solution to a system of equations as a function that assigns objects to each indeterminate and satisfies all the equations in the system.

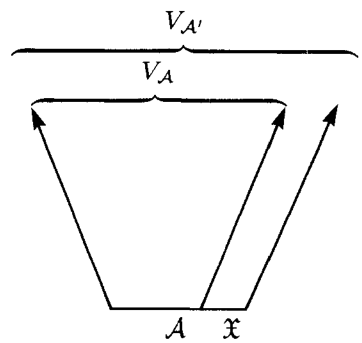

Figure 11

a unique solution in the universe of hypersets. This result, which we call the "Solution Lemma," is used repeatedly throughout the book. The remainder of this section (to page 51) is devoted to a precise formulation of this lemma, and can be skipped by anyone who finds the formulation just given precise enough.

Given a collection $\mathcal{A}$ of atoms, let us write $V_{\mathcal{A}}$ for the hyperuniverse of all sets with atoms from $\mathcal{A}$, assuming ZFC/AFA, of course. Given some larger collection $\mathcal{A}' \supseteq \mathcal{A}$ of atoms, we may also consider the hyperuniverse $V_{\mathcal{A}'}$ of all sets with atoms from $\mathcal{A}'$. Since the sets in $V_{\mathcal{A}}$ are those depicted by arbitrary graphs with tags chosen from $\mathcal{A}$, and likewise for $V_{\mathcal{A}'}$ and $\mathcal{A}'$, it is clear that $V_{\mathcal{A}} \subseteq V_{\mathcal{A}'}$. (See Figure 11.)

Let us write $\mathfrak{X} = \mathcal{A}' - \mathcal{A}$ and call the elements $\mathbf{x} \in \mathfrak{X}$ *indeterminates* over $V_{\mathcal{A}}$. Think of these indeterminates as unknowns ranging over the hyperuniverse $V_{\mathcal{A}}$. By analogy with ring theory, we write $V_{\mathcal{A}'} = V_{\mathcal{A}}[\mathfrak{X}]$. Then given any set $a \in V_{\mathcal{A}}[\mathfrak{X}]$, we can construe it as a "term" in the indeterminates that occur in its transitive closure, that is, the indeterminates in $a \cup (\bigcup a) \cup (\bigcup\bigcup a) \ldots$. By an *equation in* $\mathfrak{X}$ we mean an "expression" of the form

$$\mathbf{x} = a$$

where $\mathbf{x} \in \mathfrak{X}$ and $a \in V_{\mathcal{A}}[\mathfrak{X}] - \mathfrak{X}$. By a *system of equations in* $\mathfrak{X}$ we mean a family of equations $\{ \mathbf{x} = a_{\mathbf{x}} \mid \mathbf{x} \in \mathfrak{X}\}$, exactly one equation for each indeterminate $\mathbf{x} \in \mathfrak{X}$.

In the first of the examples above, we considered $\mathfrak{X} = \{\mathbf{x}\}$ and the system of equations was simply the single equation

$$\mathbf{x} = \{\mathbf{x}\}.$$

In the second example, we had $\mathfrak{X} = \{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ and the following three equations.

$$\begin{aligned} \mathbf{x} &= \{\text{Claire}, \text{Max}, \mathbf{y}\} \\ \mathbf{y} &= \{\text{Max}, \mathbf{x}\} \\ \mathbf{z} &= \{\mathbf{x}, \mathbf{y}\} \end{aligned}$$

In both of these examples, the sets on the right-hand side of the equations are actually wellfounded, but we could also consider equations like

$$\mathbf{x} = \langle \Omega, \mathbf{x} \rangle$$

where the nonwellfounded set $\langle \Omega, \mathbf{x} \rangle$ occurs on the right-hand side.

We next define what we mean by a solution to a family of equations, in the natural way. By an *assignment* for $\mathfrak{X}$ in $V_{\mathcal{A}}$ we mean a function $f : \mathfrak{X} \rightarrow V_{\mathcal{A}}$ which assigns an element $f(\mathbf{x})$ of $V_{\mathcal{A}}$ to each indeterminate $\mathbf{x} \in \mathfrak{X}$. Any such assignment f extends in a natural way to a function $\hat{f} : V_{\mathcal{A}}[\mathfrak{X}] \rightarrow V_{\mathcal{A}}$. Intuitively, given some $a \in V_{\mathcal{A}}[\mathfrak{X}]$ one simply replaces each $\mathbf{x} \in \mathfrak{X}$ by its value $f(\mathbf{x})$. (To make this rigorous, one has to work with a canonical graph depicting a, replacing any childless nodes tagged by an indeterminate $\mathbf{x} \in \mathfrak{X}$ with a graph depicting the set $f(\mathbf{x})$.) Rather than write $\hat{f}(a)$, we write $a[f]$, or even more informally, $a(\mathbf{x}, \mathbf{y}, \ldots)$ and $a(f(\mathbf{x}), f(\mathbf{y}), \ldots)$.

An assignment f is a *solution of an equation* $\mathbf{x} = a(\mathbf{x}, \mathbf{y}, \ldots)$ if

$$f(\mathbf{x}) = a(f(\mathbf{x}), f(\mathbf{y}), \ldots).$$

More generally, f is a *solution of a system of equations in* $\mathfrak{X}$ if it is a solution of each equation in the system.

Theorem 1 (Solution Lemma) *Every system of equations in a collection $\mathfrak{X}$ of indeterminates over $V_{\mathcal{A}}$ has a unique solution.*

This lemma is illustrated in Figure 12. Again, we stress that the lemma has two aspects, existence and uniqueness, both of which are crucial to what follows. The proof, while not difficult, is somewhat tedious, largely for notational reasons. It can be found in Aczel (1987). The following example, though, will illustrate the main idea.

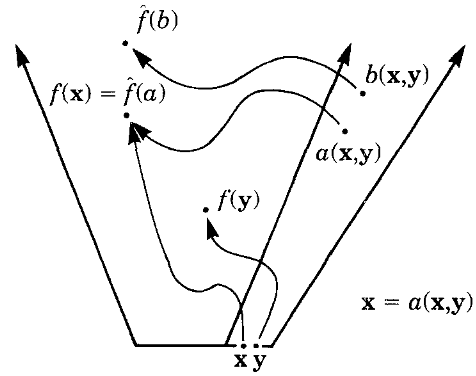

Figure 12

Example 1 Let $\mathfrak{X} = \{\mathbf{x}, \mathbf{y}\}$ consist of two indeterminates and consider the following equations.

$$\mathbf{x} = \{\Omega, \{\mathbf{x}\}\}$$
$$\mathbf{y} = \{\text{Max}, \mathbf{x}, \mathbf{y}\}$$

The sets on the right-hand side of the equations are depicted in Figure 13. To depict the solutions to the equations, we simply alter these graphs by replacing all edges terminating in a node tagged with $\mathbf{x}$ by an edge terminating in the top node of $G_{\mathbf{x}}$, and similarly for $\mathbf{y}$ and $G_{\mathbf{y}}$. This gives us the graphs in Figure 14.

By AFA, these graphs have unique decorations, and the sets assigned to the top nodes are solutions of our equations. Further-

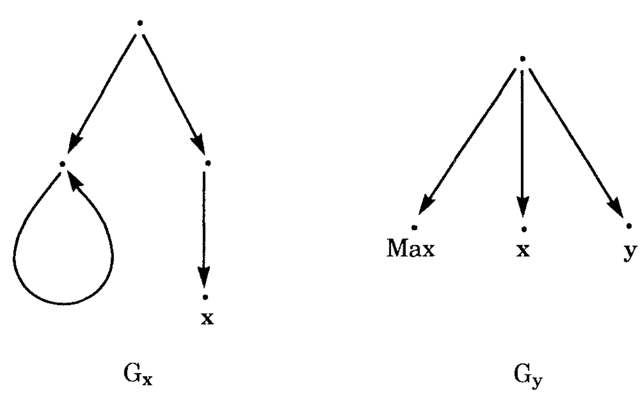

Figure 13

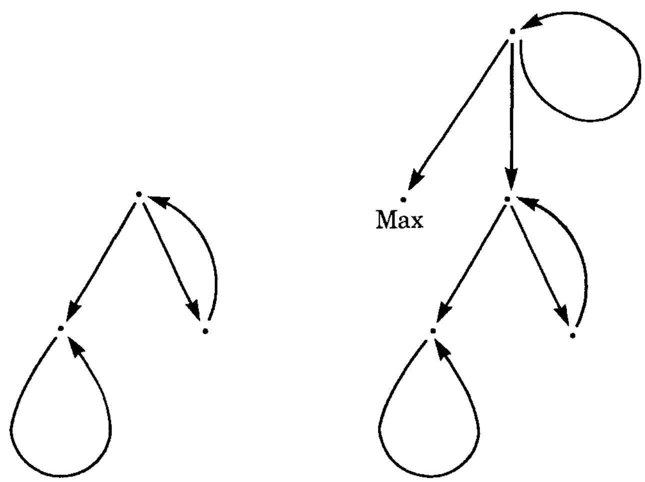

Figure 14

more, any solutions of the equations would give rise to a decoration of these graphs, so there is only one solution.

Exercise 15 Show that in the above example, the unique solution is just the assignment $f(\mathbf{x}) = \Omega$ and $f(\mathbf{y}) = a$, where a is the set depicted in Figure 15.

Exercise 16 Construct a graph depicting the set $f(\mathbf{x})$ where f is the solution of the following system of equations.

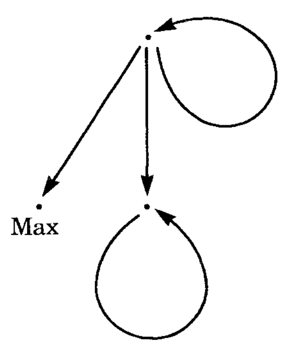

Figure 15

$$\mathbf{x} = \{\text{Claire}, \mathbf{y}\}$$
$$\mathbf{y} = \{\text{Claire}, \mathbf{z}\}$$
$$\mathbf{z} = \{\text{Max}, \mathbf{x}\}$$

Show that $f(\mathbf{x}) \neq f(\mathbf{y})$. In contrast, show that if the third equation had been

$$\mathbf{z} = \{\text{Claire}, \mathbf{x}\}$$

we would have $f(\mathbf{x}) = f(\mathbf{y}) = f(\mathbf{z})$.

Exercise 17 Working in ZFC^- (i.e., without AFA or the axiom of foundation), show that the Solution Lemma implies AFA. Thus, in the presence of the other axioms, the Solution Lemma is really a restatement of AFA.

Inductive and coinductive definitions

One final matter before we apply ZFC/AFA to model circular propositions. In set theory, a frequent technique for defining a set or class is to take the desired class to be the unique fixed point of some "monotone operator."[8] But when we work with ZFC/AFA, it often happens that there is no longer a unique fixed point, but rather many. For reasons closely connected with the Solution Lemma, it is usually the largest fixed point that is needed.

Let's look at a very simple example. Assume for simplicity that our collection $\mathcal{A}$ of atoms is finite, and consider the operator Γ that assigns to each set X the set $\Gamma(X)$ of all its finite subsets. Then if our set theory incorporates the axiom of foundation, there is a unique fixed point for this operator, the set *HF* of all hereditarily finite sets. That is, if we assume foundation, then *HF* is the unique set such that $\Gamma(X) = X$. However, in the hyperuniverse of sets, there will be many distinct fixed points, a smallest, a largest, and others in between.

The smallest fixed point HF_0 can be characterized as the smallest set satisfying the condition:

- If $a \subseteq HF_0 \cup \mathcal{A}$ and a is finite, then $a \in HF_0$.

The above is called an *inductive* definition of HF_0. By contrast, the largest fixed point HF_1 can be characterized as the largest set satisfying the converse condition:

[8] An operator Γ is *monotone* if $X \subseteq Y$ implies $\Gamma(X) \subseteq \Gamma(Y)$. X is a *fixed point* for Γ if $\Gamma(X) = X$.

- If $a \in HF_1$, then $a \subseteq HF_1 \cup \mathcal{A}$ and a is finite.

This is called a *coinductive* definition of HF_1. It is obvious from these definitions that $HF_0 \subseteq HF_1$. But in the hyperuniverse, the converse does not hold.

Exercise 18 Prove that every member of HF_0 is wellfounded. In particular, $\Omega \notin HF_0$.

Exercise 19 Prove that $\Omega \in HF_1$.

Since it seems that Ω should certainly count as a hereditarily finite set, this suggests that the coinductive definition will be the more natural one to use when working with hypersets. And indeed it is. HF_1 is just the set of those sets which can be pictured by at least one finitely branching graph. It will contain Ω and all the other examples we have given.

This is a typical phenomenon in working with hypersets. A pair of inductive and coinductive definitions which characterize the same set or class in the universe of wellfounded sets often yield distinct collections in the universe of hypersets. The smallest fixed point, specified by the inductive definition, usually consists of the wellfounded members of the largest fixed point, specified by the coinductive definition. It is usually the latter that is needed in applications.

Aczel has a theorem, the Special Final Coalgebra Theorem, which explains why coinductive definitions are so important. While the formulation of this theorem is too technical to present in detail here, we can explain the basic idea. We begin with a couple of examples to illustrate the main feature of the result.

In ZFC/AFA the Solution Lemma frequently takes the place of the Recursion Theorem of ZFC, the theorem which lets one define some operation by $\in$-recursion. To do this same sort of thing in ZFC/AFA, you show that some operation F on sets is well-defined by obtaining it as the solution to a system of equations. But then you want to know that certain properties of the equations carry over to their solutions. As long as these properties are defined by coinductive definitions, this usually works out. For example, we have the following.[9]

[9] We number theorems, propositions, and lemmas with a single numbering scheme, restarting the numbers in each of the three parts of the book.

Proposition 2 *Suppose that $\mathcal{E}$ is a finite system of equations of the form*

$$\mathbf{x} = a_{\mathbf{x}}(\mathbf{x}, \mathbf{y}, \ldots)$$

where each $a_{\mathbf{x}}$ is in HF_1. If f is the unique solution to this system, then for each indeterminate $\mathbf{x}$, $f(\mathbf{x}) \in HF_1$.

Proof: The basic idea is that if you eliminate the indeterminates from a finite set of finitary equations, in the way suggested by the proof of the Solution Lemma, you end up with a finite graph, which must then depict a set in HF_1. To do this in detail, first note that by introducing more indeterminates, we can assume each equation is of one of the following simple forms:

- $\mathbf{x} = \emptyset$,
- $\mathbf{x} = a$ (for some atom $a \in \mathcal{A}$),
- $\mathbf{x} = \{\mathbf{y}_1, \ldots, \mathbf{y}_n\}$, where the $\mathbf{y}_i$ are other indeterminates with their own equations in the system.

Let f be the solution. Since it is obvious that $HF_1 \cup rng(f)$ still satisfies the defining equation of HF_1, $rng(f) \subseteq HF_1$, as desired. □

Exercise 20 Use the above proposition to show that the unique set $a = \langle a, a \rangle$ is in HF_1.

Exercise 21 Assume the axiom of foundation, and show that $HF_0 = HF_1$. (Hint: Prove by induction on the rank of the well-founded set a that if $a \in HF_1$ then $a \in HF_0$.)

To give a second illustration of the basic notion, let's use hypersets to provide a model of what are called "streams" in computer science. The basic idea is that a stream is a possibly infinite sequence of elements. But rather than think of streams as functions from natural numbers to elements, the computer scientist thinks of them as ordered pairs, the first element of which is an atom, the second a stream. So for example, the following would be a stream.

$$\langle \text{Max}, \langle \text{Claire}, \langle \text{Max}, \langle \text{Claire}, \ldots \rangle \rangle \rangle \rangle$$

To provide an interesting illustration of the Special Final Coalgebra Theorem, let's model not just streams but arbitrary nested sequences. Given some set A atoms, let A_* be defined inductively as the smallest set containing A and closed under the rule: if $x, y \in A_*$ then $\langle x, y\rangle \in A_*$. Similarly, let A^* be defined coinductively as the largest set every member of which is either a member of A, or else of the form $\langle x, y\rangle$, where $x, y \in A^*$. We will call the members of A^* *nested sequences* on A, and the members of A_* the finite nested sequences on A.

Exercise 22 1. Show that the axiom of foundation implies that $A_* = A^*$.

2. Show that AFA implies that A_* is a proper subset of A^*. Then, show in particular that there is a nested sequence $\langle 1, \langle 2, \langle 3, \ldots\rangle\rangle\rangle$ on the set of natural numbers.
3. Show that the unique solution to the following equations yields elements of $A^* - A_*$. Give an intuitive description of those elements.

$$\mathbf{x} = \langle \text{Max}, \mathbf{y}\rangle$$
$$\mathbf{y} = \langle \text{Claire}, \mathbf{x}\rangle$$

To illustrate the Special Final Coalgebra Theorem once more, we present an analogue of Proposition 2 for nested sequences. Let $\mathfrak{X}$ be a collection of indeterminates, and consider the class of nested sequences on $A \cup \mathfrak{X}$. That is, we allow elements of $\mathfrak{X}$ as well as elements of A as basis elements in the definition. Thinking of these indeterminates as parameters, we call the nested sequences on $A \cup \mathfrak{X}$ *parametric nested sequences on A*. The Special Final Coalgebra Theorem shows that if we use parametric nested sequences in the Solution Lemma, then the resulting solutions are themselves nested sequences.

Proposition 3 Suppose $\mathcal{E}$ is a system of equations of the form

$$\mathbf{x} = a_{\mathbf{x}}(\mathbf{x}, \mathbf{y}, \ldots)$$

for $\mathbf{x} \in \mathfrak{X}$, where each $a_{\mathbf{x}}$ is a parametric nested sequence on A. Let F be the unique solution of this set of equations. Then for each $x \in \mathfrak{X}$, $F(\mathbf{x})$ is a nested sequence on A.

Exercise 23 Prove Proposition 3.

The general case of Aczel's Special Final Coalgebra Theorem goes roughly as follows. Suppose we are given some monotone operator Γ. We can use Γ to define a largest fixed point in the universe $V_{\mathcal{A}}$. Call this collection the collection of Γ-objects. However, we can also use Γ to define the largest fixed point in the universe $V_{A}[\mathfrak{X}]$, where we adjoin indeterminates. Call this the collection of parametric Γ-objects. Aczel's Theorem shows that under very general conditions on Γ, equations involving parametric Γ-objects have Γ-objects as their unique solutions. While the general formulation of Aczel's result is somewhat complicated, the proof of this consequence in any particular case is quite straightforward. We will not use the general theorem, though we will have occasion to prove special cases of it in what follows.

A final remark on Aczel's terminology, just for the curious. From the point of view of category theory, a system of equations is dual to the notion of an algebra, and hence is called a coalgebra. Final coalgebras are final in the sense of category theory, and exist under very general conditions. AFA shows that these can often be taken to be largest fixed points of monotone operators.

Exercise 24 Consider the smallest class $B_{\circ}$ containing Max as an element and closed under the rule: if $x \in B_{\circ}$ then $\{x\} \in B_{\circ}$. Similarly, define the largest class B° satisfying: if $x \in B^{\circ}$, then x is Max or $x = \{y\}$ for some $y \in B$.

1. Show that the axiom of foundation implies that $B_{\circ} = B^{\circ}$.
2. Show that $\Omega \in B^{\circ}$.
3. Formulate and prove a version of the Special Final Coalgebra Theorem for B°.

Exercise 25 Inductive definitions are used to define classes as well as sets. For example, the class of (wellfounded) ordinals can be defined inductively as the smallest class ON such that

1. $\emptyset \in ON$,
2. if $\alpha \in ON$ then $\alpha \cup \{\alpha\} \in ON$, and
3. if $a \subseteq ON$ then $(\bigcup a) \in ON$.

Give a corresponding coinductive definition of a largest fixed point ON^* and show that $\Omega \in ON^*$. Thus one might consider the set Ω a hyperordinal. However, this is a good example of a case where one would want to use the inductive definition, since the point of defining the ordinals is as representatives of well-orderings. Hyperordinals like Ω are of no use for such purposes. Formulate and prove a version of the Special Final Coalgebra Theorem for hyperordinals.

Historical Remark: The history of AFA, and other work on nonwellfounded sets, is far more complicated than we have suggested. In particular, the axiom AFA was studied independently, and earlier, by Forti and Honsell, who called it axiom X_1. Other axioms have been proposed by Finsler, Scott, and Boffa, among others. Also, the proof of consistency of ZFC/AFA is not original with Aczel, but goes back to Forti and Honsell, Gordeev, and others. The reader is invited to consult Aczel (1987) for the history of this work. We have presented it in the way we have since, to our knowledge, Aczel was the first to see that AFA could be obtained from a coherent, intuitive conception of set, rather than just being a formally consistent axiom, and to demonstrate that it is an important mathematical tool for the modeling of various kinds of real-world circularity, not just a mathematical curiosity.

The introduction of a new sort of mathematical object has always met with considerable resistance, including such now mundane objects as zero, the negative numbers, the irrationals, the imaginary numbers and infinitesimals. We realize that some set theorists feel a similar reluctance to admit hypersets as legitimate mathematical objects. While this reluctance is perhaps understandable, it is also somewhat ironic. After all, many set theorists prior to Zermelo were working with a conception which admitted circularity, as is apparent from the formulation of Russell's paradox. Furthermore, the axiom of foundation has played almost no role in mathematics outside of set theory itself. We must admit, though, that we initially shared this reluctance, having been raised within the Zermelo tradition. But our own experience has convinced us that those who take the trouble to master the techniques provided by AFA will quickly feel at home in the universe of hypersets, and find important and interesting applications.