domain of sets which satisfies these assumptions. This discovery was regarded as a paradox just because it had earlier seemed to most mathematicians that the intuitive universe of sets did satisfy the axioms.

reactions to the paradox

Russell's Paradox is just the tip of an iceberg of problematic results in naive set theory. These paradoxes resulted in a wide-ranging attempt to clarify the notion of a set, so that a consistent conception could be found to use in mathematics. There is no one single conception which has completely won out in this effort, but all do seem to agree on one thing. The problem with the naive theory is that it is too uncritical in its acceptance of "large" collections like the collection V used in the last proof. What the result shows is that there is no such set. So our axioms must be wrong. We must not be able to use just any old property in forming a set.

The father of set theory was the German mathematician Georg Cantor. His work in set theory, in the late nineteenth century, preceded Russell's discovery of Russell's paradox in the earlier twentieth century. It is thus natural to imagine that he was working with the naive, hence inconsistent view of sets. However, there is clear evidence in Cantor's writings that he was aware that unrestricted set formation was inconsistent. He discussed consistent versus inconsistent "multiplicities," and only claimed that consistent multiplicities could be treated as objects in their own right, that is, as sets. Cantor was not working within an axiomatic framework and was not at all explicit about just what properties or concepts give rise to inconsistent multiplicities. People following his lead were less aware of the pitfalls in set formation prior to Russell's discovery.

Remember

Russell found a paradox in naive set theory by considering

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

and showing that the assumption $Z \in Z$ and its negation each entails the other.

Section 15.9

Zermelo Frankel set theory zfc

The paradoxes of naive set theory show us that our intuitive notion of set is simply inconsistent. We must go back and rethink the assumptions on which the theory rests. However, in doing this rethinking, we do not want to throw out the baby with the bath water.

diagnosing the problem

Which of our two assumptions got us into trouble, Extensionality or Comprehension? If we examine the Russell Paradox closely, we see that it is actually a straightforward refutation of the Axiom of Comprehension. It shows that there is no set determined by the property of not belonging to itself. That is, the following is, on the one hand, a logical truth, but also the negation of an instance of Comprehension:

$$\neg\exists c \forall x (x \in c \leftrightarrow x \notin x)$$

The Axiom of Extensionality is not needed in the derivation of this fact. So it is the Comprehension Axiom which is the problem. In fact, back in Chapter 13, Exercise 13.52, we asked you to give a formal proof of

$$\neg\exists \mathsf{y}\, \forall \mathsf{x}\, [\mathsf{E}(\mathsf{x},\mathsf{y}) \leftrightarrow \neg \mathsf{E}(\mathsf{x},\mathsf{x})]$$

This is just the above sentence with "$\mathsf{E}(\mathsf{x},\mathsf{y})$" used instead of "$x \in y$". The proof shows that the sentence is actually a first-order validity; its validity does not depend on anything about the meaning of "$\in$." It follows that no coherent conception of set can countenance the Russell set.

But why is there no such set? It is not enough to say that the set leads us to a contradiction. We would like to understand why this is so. Various answers have been proposed to this question.

limitations of size

One popular view, going back to the famous mathematician John von Neumann, is based on a metaphor of size. The intuition is that some predicates have extensions that are "too large" to be successfully encompassed as a whole and treated as a single mathematical object. Any attempt to consider it as a completed totality is inadequate, as it always has more in it than can be in any set.

On von Neumann's view, the collection of all sets, for example, is not itself a set, because it is "too big." Similarly, on this view, the Russell collection of those sets that are not members of themselves is also not a set at all. It is too big to be a set. How do we know? Well, on the assumption that it is a set, we get a contradiction. In other words, what was a paradox in the naive theory turns into an indirect proof that Russell's collection is not a set. In Cantor's terminology, the inconsistent multiplicities are those that are somehow too large to form a whole.

How can we take this intuition and incorporate it into our theory? That is, how can we modify the Comprehension Axiom so as to allow the instances we want, but also to rule out these "large" collections?

The answer is a bit complicated. First, we modify the axiom so that we can only form subsets of previously given sets. Intuitively, if we are given a set a and a wff $P(x)$ then we may form the subset of a given by:

$$\{x \mid x \in a \wedge P(x)\}$$

The idea here is that if a is not "too large" then neither is any subset of it. Formally, we express this by the axiom

$$\forall a \exists b \forall x[x \in b \leftrightarrow (x \in a \wedge P(x))]$$

Axiom of Separation

In this form, the axiom scheme is called the Axiom of Separation. Actually, as before, we need the universal closure of this wff, so that any other free variables in $P(x)$ are universally quantified.

This clearly blocks us from thinking we can form the set of all sets. We cannot use the Axiom of Separation to prove it exists. (In fact, we will later show that we can prove it does not exist.) And indeed, it is easy to show that the resulting theory is consistent. (See Exercise 15.68.) However, this axiom is far too restrictive. It blocks some of the legitimate uses we made of the Axiom of Comprehension. For example, it blocks the proof that the union of two sets always exists. Similarly, it blocks the proof that the powerset of any set exists. If you try to prove either of these you will see that the Axiom of Separation does not give you what you need.

We can't go into the development of modern set theory very far. Instead, we will state the basic axioms and give a few remarks and exercises. The interested student should look at any standard book on modern set theory. We mention those by Enderton, Levy, and Vaught as good examples.

Zermelo-Frankel set theory ZFC

The most common form of modern set theory is known as Zermelo-Frankel set theory, also known as ZFC. ZFC set theory can be thought of what you get from naive set theory by weakening the Axiom of Comprehension to the Axiom of Separation, but then throwing back all the instances of Comprehension that seem intuitively true on von Neumann's conception of sets. That is, we must throw back in those obvious instances that got inadvertently thrown out.

In ZFC, it is assumed that we are dealing with "pure" sets, that is, there is nothing but sets in the domain of discourse. Everything else must be modeled within set theory. For example, in ZFC, we model 0 by the empty set, 1 by $\{\emptyset\}$, and so on. Here is a list of the axioms of ZFC. In stating their FOL versions, we use the abbreviations $\exists x \in y\ P$ and $\forall x \in y\ P$ for $\exists x(x \in y \wedge P)$ and $\forall x(x \in y \rightarrow P)$.

axioms of ZFC

1. Axiom of Extensionality: As above.

2. Axiom of Separation: As above.

3. Unordered Pair Axiom: For any two objects there is a set that has both as elements.

4. Union Axiom: Given any set a of sets, the union of all the members of a is also a set. That is:

$$\forall a \exists b \forall x [x \in b \leftrightarrow \exists c \in a (x \in c)]$$

5. Powerset Axiom: Every set has a powerset.

6. Axiom of Infinity: There is a set of all natural numbers.

7. Axiom of Replacement: Given any set a and any operation F that defines a unique object for each x in a, there is a set

$$\{F(x) \mid x \in a\}$$

   That is, if $\forall x \in a \exists! y P(x, y)$, then there is a set $b = \{y \mid \exists x \in a P(x, y)\}$.

8. Axiom of Choice: If f is a function with non-empty domain a and for each $x \in a$, $f(x)$ is a non-empty set then there is a function g also with domain a such that for each $x \in a$, $g(x) \in f(x)$. (The function g is called a *choice function* for f since it chooses an element of $f(x)$ for each $x \in a$.)

9. Axiom of Regularity: No set has a nonempty intersection with each of its own elements. That is:

$$\forall b[b \neq \emptyset \rightarrow \exists y \in b(y \cap b = \emptyset)]$$

Of these axioms, only the Axioms of Regularity and Choice are not direct, straightforward logical consequences of the naive theory. (Technically speaking, they are both consequences, though, since the naive theory is inconsistent. After all, everything is a consequence of inconsistent premises.)

Axiom of Choice

The Axiom of Choice (AC) has a long and somewhat convoluted history. There are many, many equivalent ways of stating it; in fact there is a whole book of statements equivalent to the axiom of choice. In the early days of set theory some authors took it for granted, others saw no reason to suppose it to be true. Nowadays it is taken for granted as being obviously true by most mathematicians. The attitude is that while there may be no way to *define* a choice function g from f, and so no way to prove one exists by means of Separation, but such functions exists none-the-less, and so are asserted to exist by this axiom. It is extremely widely used in modern mathematics. The

Axiom of Regularity or Foundation

Axiom of Regularity is so calledbecause it is intended to rule out "irregular" sets like $a = \{\{\{\ldots\}\}\}$ which is a member of itself. It is sometimes also called the Axiom of Foundation, for reasons we will discuss in a moment.

You should examine the axioms of zfc in turn to see if you think they are true, that is, that they hold on von Neumann's conception of set. Many of the axioms are readily justified on this conception. Two that are not aren't obvious are the power set axiom and the Axiom of Regularity. Let us consider these in turn, though briefly.

Sizes of infinite sets

sizes of powersets

Some philosophers have suggested that the power set of an infinite set might be too large to be considered as a completed totality. To see why, let us start by thinking about the size of the power set of finite sets. We have seen that if we start with a set b of size n, then its power set $\wp b$ has 2^n members. For example, if b has five members, then its power set has $2^5 = 32$ members. But if b has 1000 members, then its power set has 2^{1000} members, an incredibly large number indeed; larger, they say, than the number of atoms in the universe. And then we could form the power set of that, and the power set of that, gargantuan sets indeed.

sizes of infinite sets

$|\, b \,|$

But what happens if b is infinite? To address this question, one first has to figure out what exactly one means by the size of an infinite set. Cantor answered this question by giving a rigorous analysis of size that applied to all sets, finite and infinite. For any set b, the Cantorian size of b is denoted $|\, b \,|$. Informally, $|\, b \,| = |\, c \,|$ just in case the members of b and the members of c can be associated with one another in a unique fashion. More precisely, what is required is that there be a one-to-one function with domain b and range c. (The notion of a one-to-one function was defined in Exercise 50.)

For finite sets, $|\, b \,|$ behaves just as one would expect. This notion of size is somewhat subtle when it comes to infinite sets, though. It turns out that for infinite sets, a set can have the same size as some of its proper subsets. The set N of all natural numbers, for example, has the same size as the set E of even numbers; that is $|\, N \,| = |\, E \,|$. The main idea of the proof is contained in the following picture:

$$\begin{array}{ccccccc} 0 & 1 & 2 & \ldots & n & \ldots \\ \updownarrow & \updownarrow & \updownarrow & & \updownarrow & \\ 0 & 2 & 4 & \ldots & 2n & \ldots \end{array}$$

This picture shows the sense in which there are as many even integers as there are integers. (This was really the point of Exercise 15.51.) Indeed, it turns out that many sets have the same size as the set of natural numbers, including the set of all rational numbers. The set of real numbers, however, is strictly larger, as Cantor proved.

Cantor also showed that that for any set b whatsoever,

$$|\wp b| > |b|$$

questions about powerset axiom

This result is not surprising, given what we have seen for finite sets. (The proof of Proposition 12 was really extracted from Cantor's proof of this fact.) The two together do raise the question as to whether an infinite set b could be "small" but its power set "too large" to be a set. Thus the power set axiom is not as unproblematic as the other axioms in terms of Von Neumann's size metaphor. Still, it is almost universally assumed that if b can be coherently regarded as a fixed totality, so can $\wp b$. Thus the power set axiom is a full-fledged part of modern set theory.

Cumulative sets

regularity and size

If the power set axiom can be questioned on the von Neumann's conception of a set as a collection that is not too large, the Axiom of Regularity is clearly unjustified on this conception. Consider, for example, the irregular set $a = \{\{\{\ldots\}\}\}$ mentioned above, a set ruled out by the Axiom of Regularity. Notice that this set is its own singleton, $a = \{a\}$, so it has only one member. Therefore there is no reason to rule it out on the grounds of size. There might be some reason for ruling it out, but size is not one. Consequently, the Axiom of Regularity does not follow simply from the conception of sets as collections that are not too large.

cumulative conception of sets

To justify the Axiom of Regularity, one needs to augment von Neumann's size metaphor by what is known as the "cumulation" metaphor due to the logician Zermelo.

Zermelo's idea is that sets should be thought of as formed by abstract acts of collecting together previously given objects. We start with some objects that are not sets, collect sets of them, sets whose members are the objects and sets, and so on and on. Before one can form a set by this abstract act of collecting, one must already have all of its members, Zermelo suggested.

On this conception, sets come in distinct, discrete "stages," each set arising at the first stage after the stages where all of its members arise. For example, if x arises as stage 17 and y at stage 37, then $a = \{x, y\}$ would arise at stage 38. If b is constructed at some stage, then its powerset $\wp b$ will be constructed at the next stage. On Zermelo's conception, the reason there can never be a set of all sets is that as any set b arises, there is always its power set to be formed later.

The modern conception of set really combines these two ideas, von Neumann's and Zermelo's. This conception of set is as a small collection which is formed at some stage of this cumulation process. If we look back at the irregular set $a = \{\{\{\dots\}\}\}$, we see that it could never be formed in the cumulative construction because one would first have to form its member, but it is its only member.

regularity and cumulation

More generally, let us see why, on the modified modern conception, that Axiom of Regularity is true. That is, let us prove that on this conception, no set has a nonempty intersection with each of its own elements.

> **Proof:** Let a be any set. We need to show that one of the elements of a has an empty intersection with a. Among a's elements, pick any $b \in a$ that occurs earliest in the cumulation process. That is, for any other $c \in a$, b is constructed at least as early as c. We claim that $b \cap a = \emptyset$. If we can prove this, we will be done. The proof is by contradiction. Suppose that $b \cap a \neq \emptyset$ and let $c \in b \cap a$. Since $c \in b$, c has to occur earlier in the construction process than b. On the other hand, $c \in a$ and b was chosen so that there was no $c \in a$ constructed earlier than b. This contradiction concludes the proof.

One of the reasons the Axiom of Regularity is assumed is that it gives one a powerful method for proving theorems about sets "by induction." We discuss various forms of proof by induction in the next chapter. For the relation with the Axiom of Regularity, see Exercise 16.10.

Remember

1. Modern set theory replaces the naive concept of set, which is inconsistent, with a concept of set as a collection that is not too large.
2. These collections are seen as arising in stages, where a set arises only after all its members are present.
3. The axiom of comprehension of set theory is replaced by the Axiom of Separation and some of the intuitively correct consequences of the axiom of comprehension.
4. Modern set theory also contains the Axiom of Regularity, which is justified on the basis of (2).
5. All the propositions stated in this chapter—with the exception of Propositions 1 and 14—are theorems of ZFC.

Exercises

15.62 ✎ Write out the remaining axioms from above in FOL.

15.63 ✎ Use the Axioms of Separation and Extensionality to prove that if any set exists, then the empty set exists.

15.64 ✎ Try to derive the existence of the absolute Russell set from the Axiom of Separation. Where does the proof break down?

15.65 ✎⋆ Verify our claim that all of Propositions 2–13 are provable using the axioms of ZFC. (Some of the proofs are trivial in that the theorems were thrown in as axioms. Others are not trivial.)

15.66 ✎⋆ (Cantor's Theorem) Show that for any set b whatsoever, $|\wp b| \neq |b|$. [Hint: Suppose that f is a function mapping $\wp b$ one-to-one into b and then modify the proof of Proposition 12.]

15.67 ✎ (There is no universal set)

1. Verify that our proof of Proposition 12 can be carried out using the axioms of ZFC.
2. Use (1) to prove there is no universal set.

15.68 ✎ Prove that the Axiom of Separation and Extensionality are consistent. That is, find a universe of discourse in which both are clearly true. [Hint: consider the domain whose only element is the empty set.]

15.69 ✎⋆ Show that the theorem about the existence of $a \cap b$ can be proven using the Axiom of Separation, but that the theorem about the existence of $a \cup b$ cannot be so proven. [Come up with a domain of sets in which the separation axiom is true but the theorem in question is false.]

15.70 ✎ (The Union Axiom and $\cup$) Exercise 15.69 shows us that we cannot prove the existence of $a \cup b$ from the Axiom of Separation. However, the Union Axiom of ZFC is stronger than this. It says not just that $a \cup b$ exists, but that the union of any set of sets exists.

1. Show how to prove the existence of $a \cup b$ from the Union Axiom. What other axioms of ZFC do you need to use?
2. Apply the Union Axiom to show that there is no set of all singletons. [Hint: Use proof by contradiction and the fact that there is no universal set.]

15.71 ✎⋆ Prove in ZFC that for any two sets a and b, the Cartesian product $a \times b$ exists. The proof you gave in an earlier exercise will probably not work here, but the result is provable.

15.72 ✎ While $\wedge$ and $\vee$ have set-theoretic counterparts in $\cap$ and $\cup$, there is no absolute counterpart to $\neg$.

1. Use the axioms of zfc to prove that no set has an absolute complement.
2. In practice, when using set theory, this negative result is not a serious problem. We usually work relative to some domain of discourse, and form relative complements. Justify this by showing, within zfc, that for any sets a and b, there is a set $c = \{x \mid x \in a \wedge x \notin b\}$. This is called the *relative complement of* b with respect to a.

15.73 ✎⋆ Assume the Axiom of Regularity. Show that no set is a member of itself. Conclude that, if we assume Regularity, then for any set b, the Russell set for b is simply b itself.

15.74 ✎⋆ (Consequences of the Axiom of Regularity)

1. Show that if there is a sequence of sets with the following property, then the Axiom of Regularity is false:
$$\ldots \in a_{n+1} \in a_n \in \ldots \in a_2 \in a_1$$
2. Show that in zfc we can prove that there are no sets $b_1, b_2, \ldots, b_n, \ldots$, where $b_n = \{n, b_{n+1}\}$.
3. In computer science, a *stream* is defined to be an ordered pair $\langle x, y \rangle$ whose first element is an "atom" and whose second element is a stream. Show that if we work in zfc and define ordered pairs as usual, then there are no streams.

There are alternatives to the Axiom of Regularity which have been explored in recent years. We mention our own favorite, the axiom afa, due to Peter Aczel and others. The name "afa" stands for "anti-foundation axiom." Using afa you can prove that a great many sets exist with properties that contradict the Axiom of Regularity. We wrote a book, *The Liar*, in which we used afa to model and analyze the so-called Liar's Paradox (see Exercise 19.32, page 555).