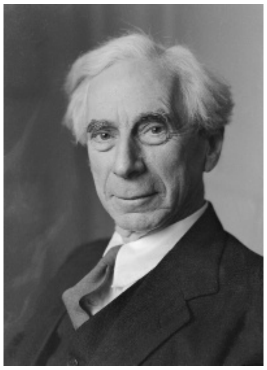
Russell's Paradox



Bettmann/Getty Images

Bertrand Russell (1872–1970)

By the beginning of the twentieth century, abstract set theory had gained such wide acceptance that a number of mathematicians were working hard to show that all of mathematics could be built upon a foundation of set theory. In the midst of this activity, the English mathematician and philosopher Bertrand Russell discovered a "paradox" (really a genuine contradiction) that seemed to shake the very core of the foundation. The paradox assumes Cantor's definition of set as "any collection into a whole of definite and separate objects of our intuition or our thought."

Russell's Paradox: Most sets are not elements of themselves. For instance, the set of all integers is not an integer and the set of all horses is not a horse. However, we can imagine the possibility of a set's being an element of itself. For instance, the set of all abstract ideas might be considered an abstract idea. If we are allowed to use any description of a property as the defining property of a set, we can let S be the set of all sets that are not elements of themselves:

$$S = \{A \mid A \text{ is a set and } A \notin A\}.$$

Is S an element of itself?

The answer is both yes and no. For suppose $S \in S$. Then S satisfies the defining property for S, and hence $S \notin S$. This contradicts the supposition that $S \in S$ and shows that $S \notin S$. Next suppose $S \notin S$. Then S is a set such that $S \notin S$ and so S satisfies the defining property for S, which implies that $S \in S$. This contradicts the supposition that $S \notin S$ and shows that $S \in S$. Thus both $S \in S$ and $S \notin S$, which is impossible because a statement is either true or false but not both. To help explain his discovery to laypeople, Russell devised a puzzle, the barber puzzle, whose solution exhibits the same logic as his paradox.

Example 6.4.3 **The Barber Puzzle**

In a certain town there is a male barber who shaves all those men, and only those men, who do not shave themselves. *Question:* Does the barber shave himself?

Solution The answer is both yes and no. If the barber shaves himself, he is a member of the class of men who shave themselves. But no member of this class is shaved by the barber, and so the barber does *not* shave himself. On the other hand, if the barber does not shave himself, he belongs to the class of men who do not shave themselves. But the barber shaves every man in this class, so the barber *does* shave himself. ■

How can the answer be both yes and no? Surely any barber either does or does not shave himself. You might try to think of circumstances that would make the paradox disappear. For instance, maybe the barber happens to have no beard and never shaves. But a condition of the puzzle is that the barber is a man who shaves *all* those men who do not shave themselves. If he does not shave, then he does not shave himself, in which case he is shaved by the barber and the contradiction is as present as ever. Other attempts at resolving the paradox by considering details of the barber's situation are similarly doomed to failure.

So let's accept the fact that the paradox has no easy resolution and see where that thought leads. Since the barber both shaves himself and doesn't shave himself, the sentence "The barber shaves himself" is both true and false. Yet the sentence arose in a natural way from a description of a situation. If the situation actually existed, then the sentence would have to be either true or false but not both. Thus we are forced to conclude that the situation described in the puzzle simply cannot exist in the world as we know it.

In a similar way, the conclusion to be drawn from Russell's paradox itself is that the object S is not a set. Because if it actually were a set, in the sense of satisfying the general properties of sets that we have been assuming, then it either would be an element of itself or not.

In the years following Russell's discovery, several ways were found to define the basic concepts of set theory so as to avoid his contradiction. The way used in this text requires that, except for the power set whose existence is guaranteed by an axiom, whenever a set is defined using a predicate as a defining property, the stipulation must also be made that the set is a subset of a known set. This method does not allow us to talk about "the set of all sets that are not elements of themselves." We can speak only of "the set of all sets that are subsets of some known set and that are not elements of themselves." When this restriction is made, Russell's paradox ceases to be contradictory. Here is what happens:

Let U be a universal set and suppose that all sets under discussion are subsets of U. Let

$$S = \{A \mid A \subseteq U \text{ and } A \notin A\}.$$

In Russell's paradox, both implications

$$S \in S \to S \notin S \quad \text{and} \quad S \notin S \to S \in S$$

are proved, and the contradictory conclusion

$$\text{both } S \in S \quad \text{and} \quad S \notin S$$

is therefore deduced. In the situation in which all sets under discussion are subsets of U, the implication $S \in S \to S \notin S$ is proved in almost the same way as it is for Russell's paradox: (Suppose $S \in S$. Then by definition of S, $S \subseteq U$ and $S \notin S$. In particular, $S \notin S$.) On the other hand, from the supposition that $S \notin S$ we can only deduce that the statement "$S \subseteq U$ and $S \notin S$" is false. By one of De Morgan's laws, this means that "$S \not\subseteq U$ or $S \in S$." Since $S \in S$ would contradict the supposition that $S \notin S$, we eliminate it and conclude that $S \not\subseteq U$. In other words, the only conclusion we can draw is that the seeming "definition" of S is faulty—in other words, S is not a set in U.

Pictorial Press Ltd./Alamy Stock Photo

Kurt Gödel (1906–1978)

Russell's discovery had a profound impact on mathematics because even though his contradiction could be made to disappear by more careful definitions, its existence caused people to wonder whether other contradictions remained. In 1931 Kurt Gödel showed that it is not possible to prove, in a mathematically rigorous way, that mathematics is free of contradictions. You might think that Gödel's result would have caused mathematicians to give up their work in despair, but that has not happened. On the contrary, there has been more mathematical activity since 1931 than in any other period in history.

The Halting Problem

Well before the actual construction of an electronic computer, Alan M. Turing (1912–1954) deduced a profound theorem about how such computers would have to work. The argument he used is similar to that in Russell's paradox. It is also related to those used by Gödel to prove his theorem and by Cantor to prove that it is impossible to write all the real numbers in an infinitely long list, even given an infinitely long period of time (see Section 7.4).

If you have some experience programming computers, you know how badly an infinite loop can tie up a computer system. It would be useful to be able to preprocess a program and its data set by running it through a checking program that determines whether execution of the given program with the given data set would result in an infinite loop. Can an algorithm for such a program be written? In other words, can an algorithm be written that will accept any algorithm X and any data set D as input and will then print "halts" or "loops forever" to indicate whether X terminates in a finite number of steps or loops forever when run with data set D? In the 1930s, Turing proved that the answer to this question is no.

Theorem 6.4.2

There is no computer algorithm that will accept any algorithm X and data set D as input and then will output "halts" or "loops forever" to indicate whether or not X terminates in a finite number of steps when X is run with data set D.

Proof (by contradiction):

Suppose there is an algorithm, CheckHalt, such that if an algorithm X and a data set D are input, then

CheckHalt(X, D) prints

"halts"	if X terminates in a finite number of steps when run with data set D

or

"loops forever"	if X does not terminate in a finite number of steps when run with data set D.

[To show that no algorithm such as CheckHalt can exist, we will deduce a contradiction.]

Observe that the sequence of characters making up an algorithm X can be regarded as a data set itself. Thus it is possible to consider running CheckHalt with input (X, X). Define a new algorithm, Test, as follows: For any input algorithm X,

Test(X)

loops forever if CheckHalt(X, X) prints "halts"

or

stops if CheckHalt(X, X) prints "loops forever".

Now run algorithm Test with input Test. If Test(Test) terminates after a finite number of steps, then the value of CheckHalt(Test, Test) is "halts" and so Test(Test) loops forever.

On the other hand, if Test(Test) does not terminate after a finite number of steps, then CheckHalt(Test, Test) prints "loops forever" and so Test(Test) terminates.

The two paragraphs above show that Test(Test) loops forever and also that it terminates. This is a contradiction. But the existence of Test follows logically from the supposition of the existence of an algorithm CheckHalt that can check any algorithm and data set for termination. *[Hence the supposition must be false, and there is no such algorithm.]*

The axioms introduced into set theory to avoid Russell's paradox are not entirely adequate to deal with the full range of recursively defined objects in computer algorithms. One response has been to develop an extension of set theory that includes new objects called hypersets. In addition, the kinds of semantic issues raised by the barber paradox are related to problems involved in processing natural language by computers.