On Russell's Paradox with Nails and Strings

Ferenc András

1985

The Russell's paradox concerns the foundations of naive set theory. This short paper is about how it can be interpreted in other contexts and has significance in the world of automata. Understanding the paper assumes that the reader is broadly familiar with the foundations of set theory and its history. The text contains many formulas and therefore the reader should be comfortable in the world of logical formulas.

In 1901, Bertrand Russell refuted the following assumption: (A) for every one-place predicate α there is a set whose members are those and only those elements for which α predicate is true. In order to refute the assumption, he employed the predicate 'not a member of itself'. Quine wrote in one of his textbooks:

> As soon as we waive his type distinctions, and read (A) simply as:[1]
>
> (A') $\exists y \forall x (x \in y \leftrightarrow \alpha(x))$,
>
> we find ourselves in trouble. For, introducing '¬①∈①' at the occurence of 'α', we can deduce a palpable falsehood as follows:
>
> (1) $\exists y \forall x [x \in y \leftrightarrow \neg(x \in x)]$
>
> (2) $\forall x [x \in y \leftrightarrow \neg(x \in x)]$ (1) y
>
> (3) $y \in y \leftrightarrow \neg(y \in y)$ (2)
>
> (4) $\exists y [y \in y \leftrightarrow \neg(y \in y)]$ (3)
>
> This difficulty is called *Russell's paradox*, for its discoverer (1901).

[1] Quine, Willard Van Orman. *Methods of Logic*. London: Routledge & Kegan Paul (1958): 249.

In order to fully grasp the assumption and the essence of the proof, I present a rather obvious example. I have slightly changed the original proof, but only to the extent that its kernel remains untouched.

It is possible to explain sets (or hypersets) in different ways, depending on the purpose of the explanation. Some formalisations of set theory can be illustrated using plane figures. Here, I reject the simple plane figure illustration, as it cannot be used to demonstrate that a set is a member of another set or of itself. Instead, I have tried to look for an illustration that sheds more light on the nature of the problem. Let us assume a room in which there are steel nails driven into the floor. Some of the nails are connected in the following way: two nails can be connected in one direction in a maximum of one way; and any string connecting two nails has its beginning and end marked. (Figure 1) The arrows represent the element relation. From the figure it can be seen that: $A = \{A\}; B = \{C, D\}; C = \{C\}; F = \{D, G, F\}; H = \{D, G\}$.
As can be seen, some of the nails are tied with a loop, others are not. We

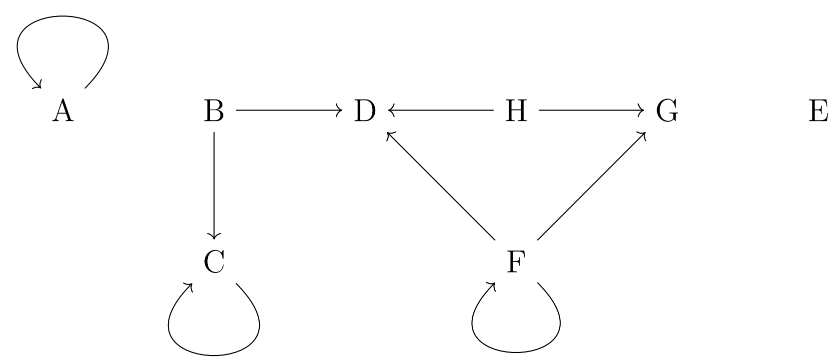

Figure 1:

then have the following task. We are given a silver nail and a brass nail and, and we are expected to connect the former with every nail in the room without a loop around it but with a string starting from it, and the latter with every nail in the room with a loop around it and with a string starting from it. The connection must always start from the brass or the silver nail. Our task is complete only if we find the unique correct solution. Here is a picture of the room again, together with the two new nails, in order to solve the task. (Figure 2)

The silver nail does not need to be connected with nails g, d and e, because they do not have a string starting from them; nor does it need to be connected with nails a, c and f, because they have a loop around them. It must, however, be connected with nails b and h, because they have strings starting from them but no loop. The question now arises: Should the silver nail be connected with itself or not? Since it has a string starting from it towards nails b and h, and has no loop around it, it should be, but as soon as we do so, we are no longer allowed to do so, because it has a loop around it. Thus, the task is unsolvable for the silver nail. In the case of the brass

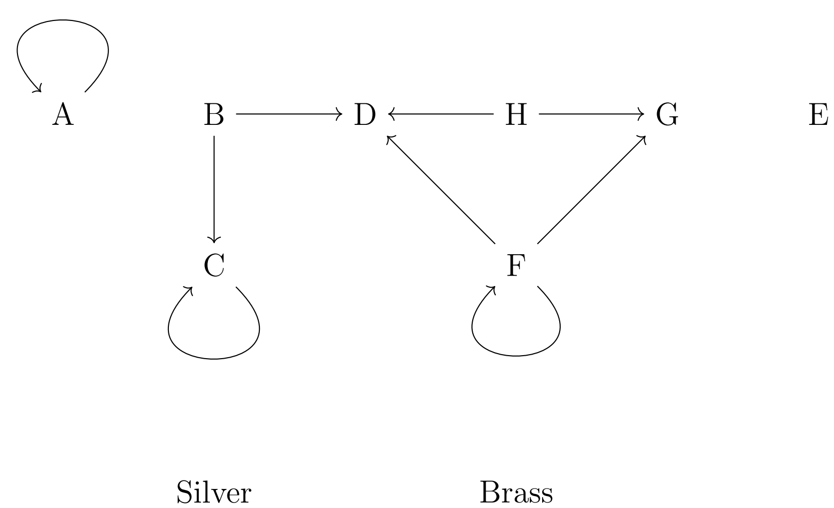

Figure 2:

nail, we find two solutions. This again is a problem, because we are at a loss as to which one to choose.

If we consider the nails as sets (hypersets) or their members, and the strings as the relation 'member of', we can formulate the task using the following two interpreted formulas:

(1)	$\forall x : x \in \text{silver} \leftrightarrow (\exists y\, y \in x \& x \notin x)$	
	$\forall x : x \in \text{brass} \leftrightarrow (\exists y\, y \in x \& x \in x)$	
	The task is unsolvable for the silver nail because (1) leads to a contradiction:	
(2)	$\text{silver} \in \text{silver} \leftrightarrow (\exists y\, y \in \text{silver} \& \text{silver} \notin \text{silver})$	(1)
(3)	$b \notin b \& d \in b)$	Figure 1
(4)	$\exists y (y \in b \& b \notin b)$	(3)
(5)	$b \in \text{silver}$	(1)(4)
(6)	$\exists y (y \in \text{silver})$	(5)
(7)	$\exists y (y \in \text{silver}) \rightarrow (\text{silver} \in \text{silver} \leftrightarrow \text{silver} \notin \text{silver})$	(2)
(8)	$\text{silver} \in \text{silver} \leftrightarrow \text{silver} \notin \text{silver}$	(6)(7)

If there were no steel nails in the room, the task could be solved for the silver nail in the following way: $\sim \exists y (y \in \text{silver})$. But even in this case, the definition of the brass nail—and the equivalent set—would be ambiguous, since a set is defined with no ambiguity by its elements. In this case, neither the statement 'brass $\in$ brass' nor the statement 'brass $\notin$ brass' contradicts premise (1). There are three ways of eluding the paradox.

I. In the first case, the silver and brass nails must be driven into the floor outside the room. In Russell's view, this should be done on the floor above, while in Zermelo's view it is sufficient to go as far as the hall. In this case, the solution to the task would be the same as if definition (1) were true only for the steel nails. Taking the steel nails to be sets, it is clear from the following that we must leave the room. Let 'U' be the set of the nails in the room:

(9)	$\forall x (x \in \text{silver} \leftrightarrow (x \in U \& \exists y\, y \in x \& x \notin x))$	
(10)	$\text{silver} \in \text{silver} \leftrightarrow (\text{silver} \in U \& \exists y (y \in \text{silver}) \& \text{silver} \notin \text{silver})$	(9)
(11)	$\text{silver} \in U (\text{silver} \in \text{silver} \leftrightarrow \text{silver} \notin \text{silver})$	(10)(6)
(12)	$\text{silver} \notin U$	(11)

Thus, neither the silver nail nor the brass nail will have a loop around it.

II. In the second case, we can impose the restriction that the silver and brass nails are not tied with a loop. This solution, however, cannot be applied to sets: Quine pointed out that the axiom equivalent to this solution (13) leads to a contradiction.[2] (13) $\forall F \exists y \forall x (y \notin y \& (x \neq y \rightarrow (F(x) \leftrightarrow x \in y)))$

[2] Quine, W.V. "On Frege's way out". *Mind* Vol. 64, No. 254 (Apr. 1955): 145-159.

III.The essence of the third solution is that the nails are not sets but signs; and definitions are always actions that take place in time. Introducing a definition means giving instructions, and to apply a definition is to fulfil the instructions. The mysterious nature of the task lies in the fact that the nails that are to be linked to the new ones are not defined by a list. If there are too many nails, this is not very practical, or may even be impossible. In such cases, we can use definitions, or rules. We formulate unambiguous conditions, and if they hold for a thing, then the rules must be applied; if they do not, the rules must not be applied. However, the rules are applicable only to those entities that are entirely determined by the conditions at the beginning. Thus, the solution to the task cannot be applied to the task itself, before one has finished solving it. That is to say, if an predicate of a thing is dependent on the solution of the task, then this predicate cannot be a part of the conditions given at the beginning. Thus, at the beginning, we have to take a photograph of the original state of the room, with no silver or brass nails. In this way, we can verify the solution subsequently.

As the nails are being considered as an illustration of sets, the following axioms concerning time and symbol usage must be fulfilled:

(14) Time is an infinite sequence of discrete moments.

(15) For every γ predicate, at any time t_1, there is a symbol, and at the following moment in time t_2, the symbol at t_2 denotes those and only those things that have γ predicate at t_1.
e.g. $\gamma(x) := \exists y(y \in_1^\star x)\ \&\ x \notin_1^\star x$ where:
$y \in_1^\star x := y$ is element of x at t_1
$y \in_2^\star x := y$ is element of x at t_2
$y \in_3^\star x := y$ is element of x at t_3
...
$x \in_2^\star$ silver $\leftrightarrow \exists y(y \in_1^\star x)\ \&\ x \notin_1^\star x$
(The temporal ordering of the modified membership relations is similar to Russell's type theory.)

(16) Sets are symbols that denote or do not denote things, irrespective of time.

In this weakened, nominalist conception, the assumption (A) refuted by Russell is nevertheless true.