

Minds, Machines and Gödel: A Reply to Mr. Lucas

Author(s): C. H. Whiteley

Source: *Philosophy*, Jan., 1962, Vol. 37, No. 139 (Jan., 1962), pp. 61-62

Published by: Cambridge University Press on behalf of Royal Institute of Philosophy

Stable URL: <https://www.jstor.org/stable/3748802>

---

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact [support@jstor.org](mailto:support@jstor.org).

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <https://about.jstor.org/terms>



and Cambridge University Press are collaborating with JSTOR to digitize, preserve and extend access to *Philosophy*

JSTOR

## MINDS, MACHINES AND GÖDEL : A REPLY TO MR LUCAS

C. H. WHITELEY

IN *PHILOSOPHY* for April 1961 Mr J. R. Lucas argues that Gödel's theorem proves that Mechanism is false. I wish to dispute this view, not because I maintain that Mechanism is true, but because I do not believe that this issue is to be settled by what looks rather like a kind of logical conjuring-trick. In my discussion I take for granted Lucas's account of Gödel's procedure, which I am not competent to criticise.

Lucas argues as follows. Gödel's theorem shows that for any consistent formal system it is possible to devise a formula which cannot be proved in that system, but which nevertheless is evidently true: viz., that formula which says, in effect, 'This formula cannot be proved in the system'. Since every machine is 'the concrete instantiation of a formal system', for every machine we can construct a formula which the machine cannot produce as true, but which a human mind can recognise to be true. Therefore there is at least one intellectual performance which a mind can achieve and a machine cannot. Therefore minds are essentially different from machines, and in one respect at least superior to them.

The trap in which the machine is caught is set by devising a formula which says of itself that it cannot be proved in the system governing the intellectual operations of the machine. This has the result that the machine cannot prove the formula without self-contradiction (for the formula says that it cannot be proved); but by virtue of this very fact that it cannot be proved, the formula states a truth. Lucas holds that the machine cannot assert this truth because it is a machine. It seems to me that the machine cannot assert this truth because its assertion of it has been made self-contradictory; and it is not a deficiency in a machine that it is unable to perform a self-contradictory task. The same is true of a human mind. It is possible to devise a formula which will trap a human mind—say, Mr Lucas's—in the same way that his application of Gödel traps the machine. Take, for instance, the formula 'This formula cannot be consistently asserted by Lucas'. It will be seen that this formula cannot be consistently asserted by Lucas; for, if he does assert it, he contradicts himself in asserting a formula which he denies that he can consistently assert. This being so, any reader of *PHILOSOPHY* other than Lucas can see that the formula states a truth, and can assert it. Does this show that our minds are unlike Lucas's, and in at least one respect superior, since there is at least one truth which we can grasp and state, and he cannot?

Certainly not. Lucas contradicts himself only if he asserts my formula itself. He can escape from the trap by stating what the formula states without using the formula: e.g. by saying 'I cannot consistently assert Whiteley's formula'. So, if the machine is to escape the Gödel trap, we must provide it with some way of asserting what the Gödel formula asserts without using the formula. This is what the observing human mind does. It is surely possible to enable a machine to do it. Can we not have a

machine which, when presented with a formula, always attempts to provide a proof of this formula from certain axioms and rules of inference; which, when it has found a proof, shows a green light, or spells out the words 'I have proved the formula', and when it has reached a deadlock or tried all the relevant combinations without success, shows a red light or spells out the words 'I cannot prove the formula'? Then, if these lights or words are treated as assertions, we have provided the machine with a way of asserting that the Gödel formula is unprovable by it, which is all that the human mind can do. Thus the Gödel theorem does not demonstrate a difference between minds and machines.

*University of Birmingham.*

## MINDS, MACHINES AND GÖDEL:

### ANOTHER REPLY TO MR LUCAS

F. H. GEORGE

I WOULD like to draw attention to the basic defect in the argument used by Mr J. R. Lucas (Minds, Machines and Gödel, *Philosophy*, July 1961, p. 112).

Mr Lucas there states that Gödel's theorem shows that any consistent formal system strong enough to produce arithmetic fails to prove, within its own structure, theorems that we, as humans ('minds'), can nevertheless see to be true. From this he argues that 'minds' can do more than machines, since machines are essentially formal systems of this same type, and subject to the limitation implied by Gödel's theorem.

This is a very brief summary of what is a more complex and interesting argument, which does indeed show that what we might call 'deduction systems' are limited by factors that do not limit human beings. Now the trouble is that this only disposes of *deductive machines* and these are really of no cybernetic interest in any case. Cybernetics has been almost wholly concerned with what are called 'Inductive Systems', or probabilistic machines that are capable of producing the axioms from which deductive operations start, and these are obviously beyond the range of being formal systems in the sense that makes Gödel's theorem applicable to them.

It may be argued that it has not yet been fully demonstrated that machines can perform inductions and can learn, think, etc., but whether this is so or not, it is quite certain that such machines are not liable to the criteria which Mr Lucas would like to impose. Or, to put the matter another way, as it was put by Professor D. M. MacKay at a recent meeting at Oxford, if you introduce the person (or mind) who goes beyond the purely deductive system, then there is no reason why mind and deductive system should not be collectively mechanised.

The point really is that in cybernetics we are not dealing with machines that are wholly specified in advance. They are self-programming or self-organising and their subsequent behaviour will depend upon the environ-