

REVIEW

H. A. P. Swart, *Artificiële intelligentie in de filosofie en psychologie*, (= *Filosofische Reeks*, nr. 1), Amsterdam, 1975, Centrale Interfaculteit Universiteit van Amsterdam, 45 pp. (Dfl. 5,-).

This publication (in Dutch) gives an overview of the arguments used against and in favour of the use of computers as a device to study human intelligence. The overview is systematic and very intelligible, but it neither gives a new insight into the question of use of computers in psychology, nor does it indicate any new, possible course of investigation in artificial intelligence itself.

The author introduces problems of philosophical and psychological nature, viewed from the angle of artificial intelligence. Swart concentrates his attention on the question if we can talk about intelligent or cognitive artifacts. This investigation has a purpose. The author wants to know if it is possible to use computers as a device to study human intelligent behaviour, and comes to the conclusion that it is only in cases in which machine behaviour can be called intelligent that it can be used to this end. Swart neither introduces new problems, nor creates new points of view for existing problems. He gives an account of an argument that started the moment the computer was taken into practical use. He merely repeats essential parts of opposite points of view and points to the fallacies they contain.

As a matter of fact he discusses the following problems :

1. The problem of consciousness. The difference between humans and computers lies in the fact that computers aren't self-conscious. This difference between men and computers is for several investigators a reason to say that computers, even fed with the appropriated programs, are not thinking and cannot be called intelligent artefacts.
2. The problems connected with the statement that human beings are the only intelligent organisms. Humans are human because they can take initiatives, because they are thinking. Thinking, being intelligent, is typical of human beings. This means that if you talk about intelligent machines, you are obliged to call them human beings.

3. The problem connected with the interpretation of the concept of a model. The author demonstrates that it is possible to construct an abstract schema that incorporates goal-directed behaviour. In accordance with such a schema one can build cognitive artefacts the behaviour of which is of a sufficiently high complexity to be called intelligent. Those artefacts are (just like human beings) models of the previous mentioned schema. One model of the schema need not to be absolutely identical to another model of the same schema.

4. The last problem is connected with the value of the Turing test. Different aspects and different implications of the test are enumerated and analysed. Criticisms of the test, due to other investigators, are treated in the same way.

After the discussion of these problems Swart comes to the conclusion that it is allowed to speak of cognitive artefacts. I very much agree with him. At the moment he introduces the use of those artefacts in psychology, he has to admit that the existing artefacts do not simulate all the intelligent behaviour human beings can develop. It would be wise then to speak of cognitive artefacts for certain forms of intelligent behaviour, restricting the abilities of the artefact in question. This would avoid starting arguments all over again if it is stated that existing cognitive artefacts cannot produce behaviour that we would classify under the heading 'invention', or 'originality'. In view of the present inexistence of intelligent machines with inventive capacities, I should like to suggest another description of the task of artificial intelligence than the one proposed in this essay. Let us first have a look at the latter.

On p. 24 Swart describes artificial intelligence as the study that tries to solve problems of the following nature : how can we program a computer to let him play chess, or to translate a text from Russian into English, or to proof logical or geometrical theorems, etc... Indeed, this describes fearly well the activity of the investigators in the field of artificial intelligence up to this moment. They try to construct heuristics that enable the machine to answer with appropriated behaviour in certain problem-situations. As a matter of fact, they already tried to construct general heuristic rules and they failed. On the same page, Swart says that if we have to answer the question 'Can an organism think?', we have to look for the adequate problem-solving behaviour of the organism in problem-situations. He mentions that those problem-situations might be new for the involved organism. But, if the problem situation is totally new, considering the fact that no general problem-solving heuristic rules have yet been discovered, a cognitive artefact will not be able to start solving the problem. Even when it

starts solving the problem, it will stop working at the moment that the time allowed is up without gaining extra or new information from the produced activity. In a new problem situation most human beings have the same difficulty to solve the problem. Most of them will do no better than the artefact if we only consider the output. But now we could lay stress on an important difference that exist between humans and cognitive artefacts. The human being has, at the moment that he cannot solve a problem, a particular cognitive behaviour that differs from the behaviour he produces at the moment he can solve the problem. A cognitive artefact has not.

If we want to work with cognitive artefacts, if we want to understand cognitive behaviour of human beings with the help of those artefacts, I should suggest another course of investigation in artificial intelligence. Instead of trying to program a computer to do well, we could try to program the computer to take steps which might lead to results that are of some use in future situations. It seems to me rather clear that human beings possess a certain cognitive behaviour that doesn't lead to immediate success, but that is strictly necessary for obtaining a solution in the future. I should advice in this connection to search for heuristic rules for reformulating problem situation descriptions, for heuristic rules for adopting knowledge to successful reformulations and for heuristic rules for ordering remaining problems.

Although Swart is aware of the remaining difficulties, he avoids to enter into them. He has to admit that the existing artefacts cannot cope with all the intelligent behaviour of human beings. He excuses this lack of completeness with the same arguments Turing used twenty five years ago, namely : there is no reason to doubt that, in due time, the computer will be able to produce any behaviour we call intelligent.

*Paula Burghgraeve
Limburgs Universitair Centrum*