Indian Philosophical Quarterly, Vol. XII, No. 3, July-Sept. 1985

MAN AND MACHINE

(Artificial Intelligence And The Philosophy Of Mind)

I

CAN MACHINES THINK? A. M. TURING'S IMITATION GAME: TURING MACHINE.

One may doubt whether God succeeded making man in his image, but not man's desire to make machines in his image. Till the philosophers bungled, the immortal divine spark, call it whatever you like, soul/mind/consciousness, as the unique differentia of man, was generally believed. Man is supposed to be a conscious being. Then man discovers he can make machines. He desires to reproduce himself unconventionally in his artifacts forcing upon nature a radical division of labour. In a sweepingly democratising process unprecedented in history, man confers citizenship upon the machine, proposing an uncanny status quo. This is a surprise, considering man's refusal to recognize more than half of his fellow beings as equals.

Let us imagine man arguing before the Creator:

Oh Lord! I am not sure if you have breathed in consciousness in me. There is no proving it logically (You know quite well that I can't accept anything sans Q.E.D.). If I argue that to be conscious is a matter of self-awareness for me I land in solipsism. Meanwhile, machines constructed on mechanical principles can simulate whatever I can perform. Now, from the observer's standpoint, there is nothing in the observed act to inform us whether it is conscious or unconscious; and we have to confine ourselves to the observable act itself. If machines can do whatever men can do, then the same principles which explain how machines function, will also explain human actions. At last we have succeeded in reproducing ourselves in a truly human way (i.e., under controlled conditions, free from the modalities of biological laws). Philosophers will

be freed from the trouble of relating consciousness to matter; existences of diametrically opposite dimensions.

We may as well imagine the Almighty Creator assuring his wayward child :

The difference be better sustained with no fear of solipsism. I, in my capacity as omniscient Being have a direct knowledge of man's consciousness whom, therefore, I know to be different from machines. And I need not infer from observable behaviour to consciousness.

Prompt is the rejoinder from men :

Lord ! The rules of the game is not fair ! You have not made us omniscient, we, who must labour within the rules of logic to avoid lurking monsters of error. Either make us omniscient (which will be the end of Philosophy anyway) or be guided by the rules of proof in ordinary discourse. If the latter is the case, you will find the situation is something like this;

Inference from the observable behaviour to consciousness is impossible because of an unbridgable logical gap.[1] Certain kinds of activity we call conscious/intelligent. Supposing machines can perform the same, shall we not call them conscious or capable of thinking? We may, of course choose to call conscious only such acts emanating from biologically reproduced beings even though these can be simulated by mechanically organised structures. But this is an arbitrary move devoid of philosophical interest. The philosophical question is rather, if the so-called conscious/intelligent acts can be reproduced by machines organised on mechanical principles, is not consciousness explained by similar principles? Or rather is it not possible to say that a machine can think, without contradicting ourselves?

To prove his point, the man proposes as imitation game, involving a man, a machine (differences in their appearance is besides the point) and an interrogator. The Almighty himself is most welcome to play the part of the interrogator so long as He is willing to submit to the rules of logic. The part of the man is to do things which the machine will not be able to, thus

helping the interrogator to identify him. That of the machine, imitating the man to escape the interrogator's identification, considering how fast digital computers are growing in storage capacity, it is theoretically quite conceivable to have in not too distant future a digital computer (Turing machine) which can play the game so well as to avoid detection by the interrogator.[2]

At this point, the Creator will be advising not to stretch the theoretical possibility beyond the Day of Judgement. He might even disappear to cut the dialogue short.

Over to Godel and earthling opponents of mechanism.

II

CAN A TURING MACHINE BE EVER CONSTRUCTED, I.E., IS IT THEORETICALLY POSSIBLE FOR A MACHINE TO DO WHATEVER A MAN DOES? MACHINES AND GODEL'S THEOREM.

We will begin with a formal property of Formal Systems proved by Godel.

For any consistent Formal System L, there is a formula which says, in effect;

"This formula is unprovable-in-the-system."

It can be shown that if this formula were provable-in-the-system, we will have a contradiction, in the following way:

If the formula "This formula is unprovable-in-the system" is provable in the system, then it would not be unprovable-in-the-system, so that "This formula is unprovable-in-the-system" would be false.

But if the formula is provable-in-the-system, then it would not be false but would be true; since in any consistent system nothing false can be proved-in-the-system but only truths.

Further, if the formula "This formula is unprovable-in-the-system" is unprovable in the system, then it is true that the formula is unprovable-in-the-system, that is, "This formula is unprovable-in-the-system" is true.[3]

Godel's Theorem has demonstrated this proof rigorously.[4]

Next is the point that a machine/a digital computer/a Turing machine is an instantiation of a formal system whose w f f are represented in numbers (Godel numbering). The initial states of the machine and its environment correspond to the axioms of the formal system, is invariant operations to the definitions, its operational procedures to the rules of inference and the results of its operation upon its input to the theorems which follow in the formal system.

Given this the argument follows:

A machine is a formal system for which there is always a Godelian formula G unprovable-in-the-system. That is to say, a consistent machine will not be able to produce its Godelian formula G as true. But we human beings can produce and see such Godelian formula as true. Here is at least one thing which no machines, but man, can do. Hence men are not machines.[5]

The above given argument is a logical version of the Liar's Paradox. Consistent lying means asserting a proposition P as true while knowing that 'P is untrue' is true. Knowing that 'P is untrue' is true cannot be made part of the assertion of the proposition P, for that will be asserting P. $\sim$ P. Hence it has to be a second order act based on the first order one. "The proposition P is untrue" is a meta-statement on P. Endowing a suspended animation to P and sustaining its negation by a higher order act is the cognitive structure of what we call lying, which again points to the negative structure of the human mind. A computer can malfunction but not tell lies, for it has no intention (read mind) to do so. If on one side, we have logical demonstration in Godel's Theorem; on the other side we have a phenomenological description in Jean-Paul Sartre's analysis of bad faith, i.e. of those cases where one deceives not others, but oneself.[6]

So much for the Godlelian arguments on the functional identification of man and machine. Some more arguments may be suggested.

Identity is a symmetrical relation. If A = B, the B = A. Hence to establish functional identity of man and machine, it is to be shown not only that machines can do whatever men do, but also men can do whatever machines do. A point of debate is whether this can be established at all. For example, can man not commit error in the likeness of a computer. "To err is human", so goes the saying. But is it part of the nature of the computer? Can a man be 'perfect' in the way a computer is? Or take for example, the bureaucracy. Perhaps the more a man reduces himself to being the machine, the nearer he is to being the perfect bureaucrat. The irreducible human factor asserting itself is perhaps what spares human civilisation from the 'perfect' bureaucracy. A computer may smoothly direct N-Bomb carrying missiles to civilian areas whereas a human pilot commissioned for the job may go insane before dropping it.

The only requirement for the eventual mechanical performance of any human act, so the mechanists may argue, is that the act be precisely explicated with reference both to the antecedent conditions upon which it ensues (the input) and to the result achieved in response to these conditions (the output). Here it may be argued that, for some kinds of human activity it is not possible to fulfil this requirement. Not that the act in question is not well understood; but a demand of explication with the input-processing principle-output schema hardly makes sense.

Take, for example, the case of religious conversion. An unknown factor effecting a complete break in a man's life is the essence of religious conversion.[7] Or a young girl falling in her first love (except in Hindi films where input-output of love can be related to a mechanical perfection).

III

ASSUMING THAT MACHINES CAN DO WHATEVER MEN DO, CAN WE SAY THAT MEN AND MACHINES PRESENT HOMOLOGOUS STRUCTURES?

The dialectics of points and counterpoints goes on, mecha-

nists (A. M. Turing,[8] Kenneth M. Sayre[9] and others) giving their rejoinders to the argument from Godel's Theorem and opponents of mechanism (J. R. Lucas[10] and others) furnishing their counterpoints. I intend to shift the ground of debate by asking the question;

Granting the functional isomorphism of men and machines as a theoretical presupposition, does it follow that they present homologous structures?

We may begin by drawing a distinction between structure and function. A structure is an organised whole that has a set of functions. Functions must be attributable either to the sructure or to the elements. Ontologically, a structure is prior to the function, as there has to be a structure to perform a set of functions. Such is the case with natural structures. But, logically, given a set of functions we may think of models of corresponding structures. Technology and social engineering furnishes examples of such.

The correspondence of function to the structure may be:

(a) A set of observable functions corresponding to homologous structures, i.e., given a set of such functions $x^1, x^2, \ldots x^n$, two homologous structures A and B perform the set. Here, not only the observable acts are identical but the mode of reproduction, essence and structural properties are also identical.

(b) A set of observable functions corresponding to non-homologous structures; i.e., given a set of such functions $x^1, x^2, \ldots x^n$, two non-homologous structures A and B perform the set. Here the observable acts are identical but the mode of their reproduction, essence and structural properties are not identical.

Given these possibilities, we now ask the moot question, how do we establish the correspondence?

There are two ways:

(a) $A \supset (x^1, x^2, \ldots x^n)$

From the structure we proceed to the set of functions. The knowledge of the structure is a sufficient condition for a know-

ledge of the range of function that will follow. For to know a structure is to know the elements in relation to the whole, the whole in relation to the elements and the principle which organises this relationship. It amounts to a comprehensive knowledge of the foundation of functionality. If we know that A and B are homologous structures and A can perform the set $x^1, x^2, \ldots x^n$, then, we can infer that B also can perform the same set.

1. $A = B$
2. $A \supset (x^1, x^2, \ldots x^n)$
3. $(A \supset B) \cdot (B \supset A)$ from 1, by → Definition
4. $B \supset A$ from 3, by → Simplification
5. $B \supset (x^1, x^2, \ldots x^n)$ from 4 and 2, by Hypothetica Syllogism.

This is an inference from homologous structures to functional identity.

(b) Correspondence from the function to the structure :

Given a set of functions $x^1, x^2, \ldots x^n$ we proceed to find the structure to which the set can correspond. In cases where the function is logically prior to the structure the correspondence is not difficult, for then, the structure is what we design and we have the following relation :

$$[(x^1, x^2, \ldots x^n) \supset B] \cdot [B \supset (x^1, x^2, \ldots x^n)]$$

Where the structure is ontologically prior to the function, establishing their correspondence is one of hypothetical approximation. In such a case, we have the relation;

$$A \supset (x^1, x^2, \ldots x^n)$$

but not the relation

$$(x^1, x^2, \ldots x^n) \supset A$$

There is structure A means it implies functions $x^1, x^2 \ldots x^n$, but there is the set of functions $x^1, x^2, \ldots x^n$ need not imply that there is structure A unless the set exhaust the ontological modalities of A. In case the set exhaust the modalities, it can be construed as logically prior to the structure and there will be no point in regarding A as ontologically prior. Our move

from the set to the structure will not be a relation of logical entailment but a hypothetical approximation that can only point to an ontological priority.

Given two structures then, A which is ontologically prior to function, and B whose function is logically prior, if we have

(1) $A \supset (x^1, x^2, \ldots x^n)$

(2) $[(x^1, x^2, \ldots x^n) \supset B] \cdot [B \supset (x^1, x^2, \ldots x^n)]$

From this we cannot infer $A = B$, unless this identity is established beforehand. That is to say, we cannot establish one-to-one correspondence between (1) and (2), for example, in the following manner:

(2) $[(x^1, x^2, \ldots x^n) \supset B] \cdot [B \supset (x^1, x^2, \ldots x^n)]$

(3) $[(x^1, x^2, \ldots x^n) \supset A] \cdot [A \supset (x^1, x^2, \ldots x^n)]$

in which case, $A = B$. But this is not possible since we cannot derive (3) from (1).

Back to man and machine. Man is a highly structured being. He is the foundation of his actions. The science of man has not been able to exhaust the ontological modalities of man, i.e., the modalities are open. Understanding man the highly structured being from his actions is a matter of hypothetical approximation.

A machine is a closed system. That is its essence, for, anything different would not be a machine. Its function set a limit to the structural modalities.

The mechanist inference that men are machines is a logical confusion between two different conditions of proof; ontologically prior structures where the knowledge of structure is a sufficient and necessary condition for understanding the mode of correspondence to the function, and structures whose functions are logically prior where a knowledge of the function is a sufficient condition for charting out all the possible structure to which correspondence can be made.

Researches on the 'artificial intelligence' are significant as a programme for exploring the possibilities of mechanical reproduction of observable part of human action. But the pheno-

menology of the act is not exhausted by the observable part alone, it includes the mode of correspondence to the structure and the organising principle. To advance the thesis that researches on 'artificial intelligence' will one day throw up mechanistic models of man is misconstrueing the nature of the programme.

D. M. College, **SOYAM LONKENDRAJIT**
IMPHAL, MANIPUR.

REFERENCES

1. Ayer, A. J., "One's Knowledge Of Other Minds", *Philosophical Essays*, (London, Macmillan and Co. Ltd., 1969), pp. 195-197.
2. Turing, A. M., "Computing Machinery And Intelligence", *Mind*, LIX, No. 236 (1950) Reprinted in
Turing, A. M., "Computing Machinery And Intelligence", *Minds And Machines*, ed. Anderson, Alan Ross (New Jersey, U.S.A. : Prentice Hall, INC, Englewood Cliffs, c 1964), pp. 4-50. Reference is to the latter source.
3. Lucas, J. R., "Minds, Machines And Godel", *Philosophy*, XXXVI (1961), pp. 112-127, Reprinted in
Lucas, J. R., "Minds, Machines And Godel", *Minds And Machines*, ed. Anderson, Alan Ross (New Jersey, U.S.A. : Prentice Hall, INC, Englewood Cliffs, c 1964), pp. 43-59, Reference is to the latter source.
4. Godel, Kurt, *Lectures At Institute Of Advanced Study*, Princeton, N. J. 1934, See also Ernest Nagel and J. K. Newman, *Godel's Proof*, London 1959.
5. Lucas, op. cit. See also, Lucas, *The Freedom Of The Will*, (Oxford : Oxford University Press, 1970), pp. 124-72.
6. Sartre, Jean-Paul, *Being And Nothingness : An Essay on Phenomenological Ontology*, trans. Barnes, Hazel E. (First Publication : Great Britain : Metthuen & Co. Ltd., London, 1957), pp. 47-67.
7. Moving instances are recorded in *The Varieties Of Religious Experience* by William James. (London : Longmans, Green and Co., 1952).
8. Turing, op. cit.
9. Sayre, Kenneth M., "Philosophy And Cybernetics", *Philosophy And Cybernetics*, ed. Crosson, Frederick J. and Sayre, Kenneth M., (New York : Simon and Schuster, 1968).
10. Lucas, op. cit.