The Third Prong:
From Descartes to Turing

(1)

IN ORDER TO be precise about the nature of the challenge that is flung across the centuries from Descartes to the technologists of our own day, four preliminary clarifications are necessary.

In the first place, it is necessary to distinguish between computers that are programmed to perform in certain ways and what I am going to call "robots"—machines built for the purpose of simulating human intelligence in its higher reaches of learning, problem-solving, discovering, deciding, etc. [1]

We can eliminate from further consideration all computers that are completely programmed. The programmed computer does only what it is programmed to do. The program that is put into it by man determines its performance—a certain output on the basis of a certain input. It can be programmed only for performances that are logical. Its chief superiority to man lies in its speed and its relative freedom from error. Its chief utility is in serving man by extending his power, just as a telescope or a microscope does. [2]

Robots in principle are different from programmed computers. Instead of operating on the basis of predetermined pathways laid

down by programming, they operate through flexible and random connections. They may have what Turing calls "infant programming"—some fixed connections that are like those built into the human infant at birth. But for the purpose of simulating human intelligence in its higher reaches, the robot must be able to learn from its own experience and must be teachable, as the human being is. Like human beings and unlike computers, robots must be capable of making errors that are not wholly explicable by mechanical defects, as are the errors of a programmed computer.

In the second place, we must distinguish between simulation and what is called "replication." The attempts to construct mechanical models that operate in the same way that the human brain operates are efforts at replication. Most of these efforts up to the present have not gone beyond the stage of mathematical theory and the drafting board; a few actual models, such as the McCulloch-Pitts nerve net, have been constructed. These attempts to replicate the action of the central nervous system have been criticized by leading neurologists on two grounds: first, on the ground tht we do not yet know enough about the action of the CNS to attempt its mechanical replication; and second, on the ground that all such efforts are severely limited by the crucial difference between the electrochemical action of the CNS and the purely electrical action of the mechanical models. [3]

In contrast to replication, the simulation of human behavior by machines consists in achieving the same end result in the way of performance but not achieving it in the same way. Thus, for example, airplanes simulate the flight of birds, but the mechanics of flight are not the same in both cases, though both bird and airplane obey the same laws of aerodynamics. Machines have been built to simulate trial-and-error learning, pattern-recognition, chess and checker playing, working with hypotheses, carrying out a systematic search for solutions to a problem; but in all these cases the machine—whether a programmed computer or something more like a robot—achieves the result without the working of its machinery replicating the neurophysiological processes of the human being who accomplishes the same result. [4]

We can ignore replication and concern ourselves only with simulation; for the Cartesian challenge only calls for a machine

that can simulate conceptual thought as that is exhibited in the flexible and unpredictable give-and-take of human conversation. It is possible to enumerate a whole series of differences between machines and men, beginning with the fact that machines are not alive, but though such differences may be significant for the problem of replication, they do affect the problem of simulation. In short, if a machine were to engage in conversation with men, using an ordinary language such as English and using it as men use it, the Cartesian challenge would be met, even if the machine otherwise differed from men in many and various respects.

In the third place, we must distinguish, in this field of technology, between actual achievements and theoretical promises for the future. As might be expected, the claims that are made both with respect to mechanical devices now actually in operation and also with respect to the theoretical possibilities that will be realized in the future, range from the very modest at one extreme to the most extravagant at the other. [5] The explanation of the deficiencies in existing apparatus, where they are admitted, tends to support the prediction that the machines of the future will come much closer to simulating human intelligence than any now in existence. [6]

The magnitude of the human brain in componentry (number of neurons) is a very large number—from 10^{10} to 10^{11}. And the magnitude of the circuitry (number of connections) is very much larger. The magnitude of existing machines is very much smaller —both in componentry (number of transistors—10^3) and in circuitry. Until machines are built that more closely approximate the magnitude of the human brain—as, for example, the magnitude of the dolphin's brain does—it is unreasonable to expect an unprogrammed robot to simulate the most characteristic of human performances, such as learning a language like English and using it conversationally. The only further point that need be made here is that there is no reason in principle to deny the possibility of building a robot that will have a componentry and a circuitry equal in magnitude to, or even greater than, that of the human brain.

No one, not even the most extravagant among the technologists, claims that a robot now exists that can meet the Cartesian chal-

lenge; though, as we shall see presently, there are many theoreticians in this field who do not hesitate to predict that one will be produced in the future.

In the fourth place, and finally, we must carefully note that the challenge to future technologists is very precise. It does not call for the production of a robot whose performance will provide us with an answer to the loose and unclear question, "Can machines think?" The literature concerned with computers and robots quite properly dismisses that question as so loose and ambiguous that there is no way of deciding what the correct answer is. In its use by psychologists, neurologists, computer technologists, and philosophers, the word "think" has so many meanings in its application to animals, men, and machines, that if anyone asks, "Can animals think?" or "Can machines think?" the answer must be "Yes"—*in some senses of the word*, and probably also "No"—in other senses of the word. No one has yet produced an acceptable definition of human thinking *in all its variety* that will serve as a standard for measuring the success of efforts to produce a robot that will simulate the whole range of human thinking. [7]

Fortunately, in order to make a critical test of artificial or machine intelligence, it is not necessary to do what it might always be impossible to do; namely, *gain general acceptance for a definition of human thinking in all its variety*. The Cartesian challenge to the technologists calls for a very specific performance that would sufficiently indicate that the robot had the power of conceptual thought, by virtue of the fact that the robot could use propositional language conversationally.

What justifies our inference that men have the power of conceptual thought and that other animals lack this power? It is the fact that men have and animals lack propositional speech. This being so, we would be equally justified in attributing the power of conceptual thought to an unprogrammed robot that was able to engage in conversation in English; and by the same token, we would be justified in saying that a robot failing this test did not have the power of conceptual thought, no matter what other intelligent or apparently "thinking" behavior the robot manifested.

(2)

In the light of the foregoing clarifications, and with this prong of the Cartesian challenge made as sharp as possible, we are now in a position to appreciate the remarkable fact that an English mathematical genius just recently deceased, A. M. Turing, has picked up the gauntlet thrown down by Descartes more than three centuries ago. Turing's widespread fame in the world of computers and robots—and the currency of such phrases as "Turing's game" and "Turing's machine"—centers on his claim that it is mathematically possible to conceive a robot that will successfully meet Descartes' challenge.

"Turing's machine" is a mathematically conceivable robot of the future that will be able to play Turing's game as well—or almost as well—as men can play it. "Turing's game" is a conversational affair using an ordinary language, such as English. It is derived from a game in which all the players are human beings. Two of the players are behind a screen; one of them is a male, the other female. The third player is the interrogator who asks the hidden participants questions in an effort to determine which one is male, and which is female. The questions (*unlimited as to content or variety*) are submitted in typewritten form and answers return in typewritten form, so that tone of voice is eliminated as a clue. The hidden players are *not required to tell the truth* in answering. They can say anything that they think will serve to prevent their being detected. The Turing version of this game simply substitutes a robot for one of the human beings. All the rules of the game remain the same, but the problem becomes one of determining which hidden participant is a human being, and which a robot. [8]

If he were confronted with Turing's game, Descartes would say that no machine could ever be built that would be able to participate in it at all, much less be able to play it as effectively as a human being. Turing's claim, on the contrary, is that a robot participant in his game is now theoretically conceivable and that, with the development of technology, his theoretical model can some day be actually produced in the form of a machine that will use an ordinary language, such as English, with the con-

versational flexibility and resourcefulness required to play the game effectively.

Among competent scientists and technologists, there are many differences of opinion about Turing's claim, ranging from dismissal of it as unfounded to endorsement of it as sound. For reasons already given, we can ignore the type of criticism which says that human thinking involves much more than the ability to play Turing's game effectively. It may, indeed; but playing Turing's game would satisfy our criterion for attributing to the robot that played it the power of conceptual thought. We can also ignore those who point out the numerous difficulties that have so far been encountered in programming computers to use an ordinary language such as English; or those who go further and maintain that such programming is inherently impossible. That, too, may be the case; but Turing's machine will not be a programmed computer; it will be a robot with no more than infant programming, a robot able to learn English and learn how to use it conversationally. [9]

One further comment: the question is not whether Turing's mathematical theorizing is sound. It is rather whether the robot that he envisages will ever come into existence and operate as intended. *Only if and when it does will the third prong of the Cartesian challenge be successfully met.*

(3)

Let us for the moment suppose that Turing's claim is validated at some future date. *Would the significance of this fact be entirely clear, or would there be some ambivalence about it?*

In the case of the dolphins, we observed that success in engaging them in conversation could be interpreted in two ways. We would have to admit that they had the power of conceptual thought, for the same reason that we attribute it to man. But, as we saw, experimental success with the dolphins would not be decisive with regard to the question whether the brain, or a certain magnitude of brain, was the sufficient, or only a necessary, condition of conceptual thought.

Can there be more than one interpretation of success in pro-

ducing Turing's machine? *Theoretically, the answer is affirmative.*
On the one hand, we could say that an immaterial factor is present in the machine as well as in man, in view of the fact that both have the power of conceptual thought, and conceptual thought involves an immaterial factor (at least according to the direct argument expounded in Chapter 12). On the other hand, we could say that since the machine is entirely material in its constitution, and since the machine exhibits the power of conceptual thought, no immaterial factor is required, either for the robot's performance or for man's.

But while both answers are logically possible, the first can be dismissed as fanciful rather than serious, for it involves an invocation of ghosts or pixies. After all, we built the machine step by step, and every component that entered into its construction was a material component. To give the first answer seriously, we would have to claim—not seriously, I hope—that while we weren't looking, an immaterial factor crept in and hid itself in the works.

We are thus left only with the second interpretation of success in the production of Turing's machine; and, according to that interpretation, no immaterial factor is needed to explain the possession—by a machine or by a man—of the power of conceptual thought as exhibited in the indefinitely variable and, therefore, unpredictable turns of human conversation. The Cartesian challenge would be satisfactorily met. More than that, the conclusion of the direct argument for the immateriality of conceptual thought would be falsified by observable facts, and we would be obligated to re-examine the premises and the reasoning to discover the source of the error. [10]