The Mimic Minds ?

Ranjit Nair & Shobhit Mahajan

TIMES OF INDIA, JUNE 2, 1991

"There is no principle of science or engineering that prevents us from making intelligent computers that are infinitely smarter than ourselves They will talk to us only to amuse themselves and so, in some sense, keep us as pets ". This prophecy was made some years ago by a high priest of Artificial Intelligence (AI), Ed Fredkin of MIT. The sinister implications of his vision did not worry Fredkin, a passionate believer in the " effecting of all things possible", a mandate contained in the founding charter of modern science drawn up by the English philosopher Francis Bacon in the seventeenth century.

The new "thinking cap" -- to use the historian Herbert Butterfield's evocative phrase -- that mankind wore with the advent of modern science, has radically transformed the ways in which we think about the world and our place in it. The process of 'disenchanting ' the world, of exorcising nature of the anthropomorphic powers attributed to it by traditional thought, is now so near complete that we seem perilously close to disenchanting our `selves'. The `soul' or the `self', long the preserve of religious eschatologies that thrived on their proclaimed superintendence, seems on the verge of being dismissed as a fable.

The threat is not new; in the eighteenth century Enlightenment sparked off by the spectacular success of Newtonian physics, La Mettrie spoke of "L' Homme Machine." The late twentieth century version is "L' Homme Computer ".

The dominant metaphor of the computer applied to the human mind draws its strength not so much from concrete explanatory success as from, to quote the neurophilosopher Patricia Churchland, " the computer's status as 'the' Technological Marvel of our time". The metaphor in turn has spawned a myth, that of the super-intelligent machine retailed by AI gurus like Fredkin. Advocates of the `strong AI' program according to which the mind can be modeled by a computer, can point to `Deep Thought', a computer programmed to play chess , which shared a title with Grandmaster Tony Miles in a chess tournament a few years ago, beating a GM en route to victory. Chess playing, if taken as representative of intelligent behavior, cannot separate man from computerkind. If anything, humans are a lot slower than contemporary devices at computational tasks. Alan Turing proposed, some four decades ago, a test of whether an intelligent machine actually `thinks'. According to Turing's criterion, given a person and a machine and a set of questions to be put to each, if one could not tell the two apart on the basis of their answers alone, then one should ascribe thought to both in equal measure. Oddly enough, as

Roger Penrose ◎notes in 'The Emperor's New Mind : Concerning Computers, Minds and ◎the Laws of Physics, *◎ the Turing test can be nontrivially applied only by being somewhat unfair to the computers. Any question that involves rapid and complex computations would have to ◎be avoided as these would immediately tell apart the quicksilver ◎computer from the halting human.◎

What if technology advanced to the point where computers could ◎not be told apart from humans on the basis of the Turing test? ◎This may indeed happen sooner than we expect. Will ◎homo sapiens◎ ◎then be divested of their alleged unique status as thinking beings? There is a long tradition in philosophy which maintains ◎that there is a difference in kind, rather than one of degree, ◎between human minds and natural objects. Descartes, at the beginning of modern science, proposed a dualism of matter and mind as ◎two independent substances. While crediting human beings with ◎minds, Descartes was inclined to view other animals as little ◎more than complex machines. This was a concession to modern ◎science, a demarcation of territory which said `Put as much as ◎you like into the physical world, as long as you leave man's ◎vital essence - his soul - outside of it '. This Cartesian doctrine has new defenders in the face of the inroads made by artificial intelligence.◎

John Searle, a philosopher at University of California, Berkeley, has produced a "Chinese Room" argument designed to show that the ◎ability to manipulate symbols is not the same as 'understanding' ◎them in the manner of a mind. At its most powerful, the argument ◎is as follows. Consider a large number of people, say the population of India, enclosed in a room (the Black Hole of Berkeley?!). ◎Assume they do not know any Chinese or exclude those who know the ◎language. From the outside, pass them a string of Chinese symbols representing stories, a set of questions about the stories and an ◎instruction booklet, say. in Hindi (assuming again that all ◎Indians in question know Hindi) containing the rules for manipulating the Chinese symbols. In response to the questions in ◎Chinese, they are required to produce yes/no answers in Chinese. ◎

◎Searle argues that although this unfortunate population of India ◎will produce the correct answers indistinguishably from any ◎speaker of Chinese, they cannot claim to have understood the ◎stories. Manipulation of symbols through rule-governed procedures, so Searle claims, cannot constitute understanding. In ◎Searle's view, the mind is not a computer. The representational ◎nature of thought, Searle believes, cannot be reproduced by computing machinery. Unlike a machine, the mind can entertain ◎thoughts not only about the here and now, but also about the ◎past, future possibilities, and even impossible things such as ◎perpetual motion machines.◎

Other critics of the strong AI program, like Penrose, refuse to ◎concede that it is even possible for a computer to successfully ◎mimic the mind. These critics say that features

like the 'oneness' of consciousness cannot be simulated by a computer, whether ⊚it does calculations sequentially, or in parallel, as in the new ⊚generation of computers. For Penrose, the conquest of science's ⊚last frontier, the human mind, may involve the discovery of new ⊚laws of physics, or at least, contact with the quixotic realms ⊚governed by the laws of quantum physics. These proposals, however, seem consistent with the vision of ⊚homo computer⊚ , updated to ⊚take account of new physics. However, the point at which Penrose demurs is when he maintains that the creative insight, unlike ⊚routine mental activity, cannot possibly b the result of following of a set of rules as in a computer algorithm. For Penrose, ⊚creative insight of the sort involved in 'sensing' of a mathematical theorem or a musical composition as a whole, cannot be ⊚algorithmic. He uses a famous result of Godel to argue that there ⊚are mathematical truths we can sense for which no algorithmic ⊚proof can be given. Creative insight, for Penrose, is like direct access to a Platonic realm of mathematical forms. Presumably, insight of aesthetic or moral variety would also qualify as a ⊚'supersensing'. ⊚

Whether or not the bizarre prophecies of the AI enthusiasts will ⊚come to pass is hard to say. But there is a new dialogue between ⊚AI and the branch of biology that deals with the working of the ⊚brain --- neurobiology, born of realisation that each discipline has much to tell the other. There are deep implications, both for ⊚ethics and our worldview, stemming from the new models of mind. ⊚

⊚The philosopher Wilfrid Sellars once contrasted the "manifest ⊚image" of reality offered by commonsense perceptions with the ⊚"scientific image" held out by science. The advocates of strong ⊚AI seek to replace the manifest image of the mind with the scientific image of a brain-computer. However, a philosophic examination of science reveals that the scientific image is critically ⊚dependent on the manifest image despite being often at odds with ⊚it. Hence fears that mind will be (or has already been) supplanted by the brain-computer seem greatly exaggerated.⊚

* Oxford University Press 1989, Vintage paperback edition 1990.

Ranjit Nair is a philosopher of science at the National Institute of Science, Technology and Development, New Delhi.⊚

Shobhit Mahajan is a physicist at St. Stephen's College, Delhi