

The Thinking Artifice: AI and its Discontents



MODERN man is incredibly lonely and miserable in his godless universe. Having lost his faith in the existence of the Supreme Intelligence as the Prime Mover of nature and cosmos, he is desperately seeking new venues of self-discovery. Though the thought of having been created in God's image leaves him cold today, modern man is not at all averse to the prospects of seeing himself reflected in the mirror of his own creation - the thinking artifice. The search for AI, of man-made forms of 'cognition' and 'intelligence', thus forms the terminal stage of the pathology of the secular man. Ostensibly planning to obliterate all marks of distinction between his own self-transcending consciousness and the inert world of inorganic matter, man today is in fact making his final bid for the discovery of meaning to his existence. With the ultimate stake of his own identity, winning or losing everything with a single bet as it were, he is desperately seeking a certain answer to his uncertain self.

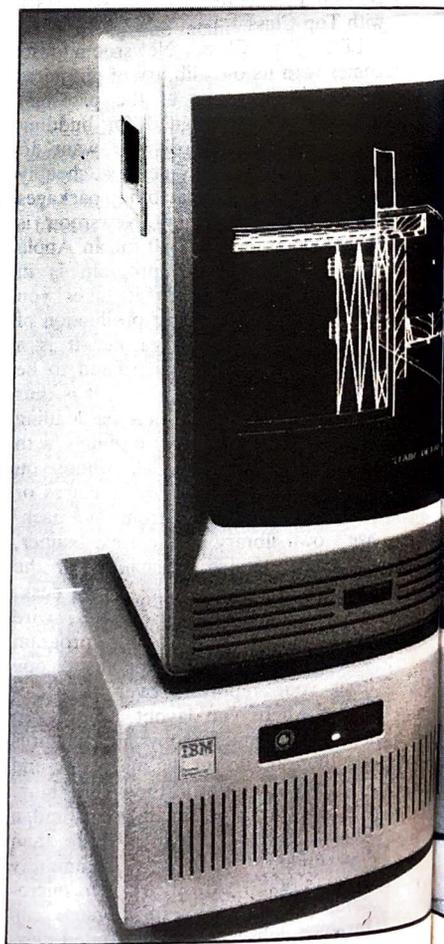
If meaning does not 'descend' from above, if there is no 'transcendence' so to speak, could there still not be 'salvation' below? If the heavens above are silent, must the matter below also remain mute for ever? Furthermore, if man is absolutely unique - and unbearably lonely - in this boundless universe, the secularist discourse continues, why shouldn't the mute world of matter be made to communicate to man, on his terms and in his own language? Why shouldn't indeed the dead stones and metals of the earth be compelled to speak at man's commands? If man the creature finds himself without a creator, the rationale against loneliness demands, let him become the creator instead and initiate his own discourse with the speechless universe. AI-speech is thus part of man's search for an answer to the silence of the Heavens and the muteness of the uni-

The quest for 'artificial intelligence' (AI) betokens modern man's restless longing to find partners in the silent universe. By striving to become a creator in his own right and initiate an 'intelligible' dialogue with the mute world of lifeless matter in his own language, the secular man also hopes to unravel the mystery of his own being. The latest Promethean bid at self-glorification, however, is foredoomed to be an exercise in futility. For, if there be no insuperable metaphysical barrier separating the world of mind from that of machines, will not the Great Nirvana of AI result in the extinction of all our mental activities, asks **S Parvez Manzoor?**

verse. Not surprisingly, then, the prospects of engendering AI make modern man finally exclaim: 'Aha! there may be a meaning to my existence after all!' Could AI, however, really take for the modern man the place of the holy communion through which the religious man found companionship with the world? For some, the answer is decidedly in the affirmative. AI for them is man's own sacrament of his future salvation! Partaking of AI would not only transform man into the likeness of God (which for the modern man never existed in the first place!) but by AI will man also be able to breathe his own spirit into the 'unintelligent' universe!

Perhaps, to view AI from the philosopher's lofty perspective is to perceive it dimly. Artificial intelligence, in its present pedestrian state at least, is a far cry from the noble image of AI-man. In fact, it has all the attributes of the deranged monster of Frankenstein - and some uncomlier details as well. At best, AI is an infantile prodigy that might prove to be an adjunct to, and not a substitution for, the human brain. At worst, it might grow to be the giant of a moron, all binary muscles and no real brain. At any rate, in its present form AI presents itself as more of a technology than a science, a toy rather than a machine, a checker player rather than a chess master. And, as with everything created by man, it is a mixture of good and bad, fraught with both promises and perils, whose ultimate answer is shrouded in the mystery of man himself. Most certainly, AI will neither supply the ultimate answer to the riddle of man, nor replace him as the main actor in this

'silent' universe. Far from controlling its human creator - against whose perversity and ignorance, however, there is no guarantee - AI will provide man with yet another tool by which his moral sensitivity will be tested. Thus,



the ultimate issue in AI, philosophy or no philosophy, is man. What does he want of it, what can he achieve by it, and how may it help him achieve self-transcendence are the central issues raised by AI. Hence, even if its chief contemporary manifestations, such as expert systems, speech recognition, robotics, vision, CAD and front-end systems, speak to us today in depressingly utilitarian tones, AI is undeniably an issue of man.

Paradoxically, though man has always prided in having intelligence and considered it his mark of distinction and honour, he has not been able to define it with any degree of certitude or unanimity. All definitions of 'intelligence' are in fact time-bound and culture-based. Sampling a few from among the thinkers of the West, in some ways foreshadowing the theoretical assumptions of AI, we come across some, for instance, that claim that intelligence is 'man's main characteristic' that it is another aspect of 'human freedom', or that its main characteristic is 'unpredictability', or, more convincingly, intelligence is the

'problem-solving ability' of man. (Here, in fact, we have the two paradigmatic views of intelligence that may be aptly described as the 'humanistic' and the 'technical!') Given our uncertainty about the nature and character of human intelligence itself (at times also described as 'reason' and 'intellect') it is not surprising that the not on as to what constitutes AI is also shrouded in the subjective vagaries of history. Moreover, as every discussion on AI inevitably gets bogged down in the man-machine quagmire, popular perceptions of man-made intelligence are always shifting grounds so as to constantly reassure us of the uniqueness of man. Once some mental process gets 'programmed' and is able to simulate human behaviour, it is no longer considered essential to 'real thinking'. The inner core of intelligence is always the unprogrammed, perhaps unprogrammable, next step. Thus, in the face of this 'humanistic subterfuge' as it were, Larry Tesler was forced to propose a theorem that claimed rather in a spirit of desperate sarcasm that 'AI is whatever hasn't been done yet'.

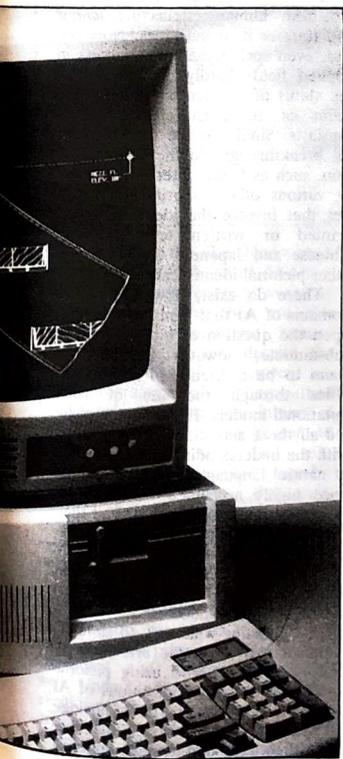
In order to narrate the story of AI, one would thus have to start at some arbitrary point. Perhaps, as good a starting point as any in this narration would be the human personality of the mathematical genius, Norbert Wiener. Despite his deep faith and concern for everything human, it is to the credit - or discredit - of Wiener to have blurred the borderline between mechanisms and organisms, or machines and men if you prefer, and made possible the development of thinking artifices. Wiener's seminal insights into the nature of life as *communication*, information exchange rather than energy transfer, laid the theoretical foundations of the science of *cybernetics* and paved the way for the appearance of machines that exchange information, give and receive feed-back, are self-regulating, and thus have acquired 'human' characteristics today.

Wiener, however, was a sensitive humanist and the prospects of 'synthetic life' and 'artificial intelligence' worried him a lot. One of his famous stories was about the rabbi of Prague, who moulded a statue of clay and by inscribing the unspoken name of the Almighty, brought it to life. The rabbi also gave it an identity by naming him as Joseph Golem. Despite Golem's considerable success as an artificial spy, the task for which he was 'created' in the first place, it became later apparent to all concerned that the

synthetic-intelligence of Golem had serious flaws. Golem often behaved erratically and menacingly and the rabbi had to destroy it by erasing the inscribed name of God on its forehead. The Goeml legend, which shows unmistakable reverberations of the Biblical theme of the Creation, provided Wiener with one of his most sombre insights. He came to believe that should humans ever succeed in creating an artificial mind, they were likely to have no more 'success' than the God of the Bible (whose pride of creation, the human Adam, let it be recalled, rebelled against his Creator and brought 'evil' to this world by his act). Wiener became fully convinced towards the later part of his life that 'humans were not ready to play god' and that any experiment with synthetic intelligence might end in disaster for both men and machines.

The next notable actor in our story of the AI was also a man of exceptional talents and sombre visions. Alan Turing, born in 1912, was the son of an Indian Colonial Service employee who grew up in England and died at the early age of 41 in 1954, possibly through suicide. Like Wiener, he was a gifted mathematician and the West's leading expert in computer sciences of his times. Turing is accredited with the distinction of extending the mind-machine similarity a point further than Wiener. Whereas Wiener regarded the boundary between organisms and mechanisms to be arbitrary, Turing claimed that mental processes can be computerised and, in that sense, there is no essential difference between a computer and a brain. Turing's claim can be understood in a variety of ways and is capable of being expressed in a number of fashions. The reductionist version of it, for instance, may be paraphrased in the words of Douglas R. Hofstadter, as that 'all brain processes are derived from a computerisable substrate'. A more accessible way of describing Turing's contribution to AI would be to say that he enunciated a functional and 'common-sense' definition of 'intelligence'. The famous Turing Test may quite simply be stated: 'If a machine can convince humans that its behaviour is intelligent, then its behaviour really is intelligent!' It is not accidental then that Turing dismissed the question of whether machines could think or not as being 'too meaningless to deserve discussion' because our definitions of *thinking* and *thought* struck him as highly arbitrary and artificial.

Turing's analogy between brains



and computers no doubt removed many covert 'metaphysical' hurdles that stood in the way of AI theorists and programmers and had hampered their progress because of the overall preoccupation with the unrewarding task of defining the 'nature' of intelligence. Intelligence is what humans recognise it as such, was Turing's answer. It is obvious that by escrowing the slippery question of the ontological definition of intelligence as some mysterious, 'transcendent' entity and by proposing a functional definition that was more in the nature of Kuhnian sociology of knowledge (Turing's work, however, preceded Kuhn's theoretical formulations), Turing was able to put all researches on AI on a solid empirical footing. His goal was to build an 'artificial brain' in which the spark of 'intelligence' would be stored so to speak, as a programmed 'software'. It ought to be noted that Turing's vision dominates many areas of research into AI to this day and many programmers, in treating intelligence as simulable human behaviour, follow his lead unquestioningly.

No matter how one assesses Turing's contribution towards the growth of AI, his theoretical uncertainty - he merely *eschewed* the question of the nature of intelligence and did not *solve* it - has clearly left its mark on all subsequent activity in this field. The most noticeable effect of this is that AI has evolved into a form of technology rather than a science. All AI research, moreover, is plagued with the absence of 'an appropriate AI methodology'. No rigorous internal criteria for self-evaluation exist today (Turing's thesis of implementability is quite vacuous), nor any well-defined notions of 'knowledge' inform its theoretical worldview. Turing's analogy between brains and computers, furthermore, is highly questionable. Though our knowledge of the functioning of the brain has increased manifold since Turing's times, we are still in no position to claim any analogical relationship between the two unknowns. (That this analogy can be carried too far is evidenced by the much over-rated work by Douglas R. Hofstadter: *Goedel, Escher, Bach*; Penguin Books, 1979. Accepting hardware/software-brain/mind analogy, he is forced to concede that though the thought of actually duplicating or simulating the hardware of the brain is 'quite aborting', still 'if intelligence involves learning, creativity, emotional responses, a sense of beauty, a sense of self, then there is a long road ahead, and it may be that these will

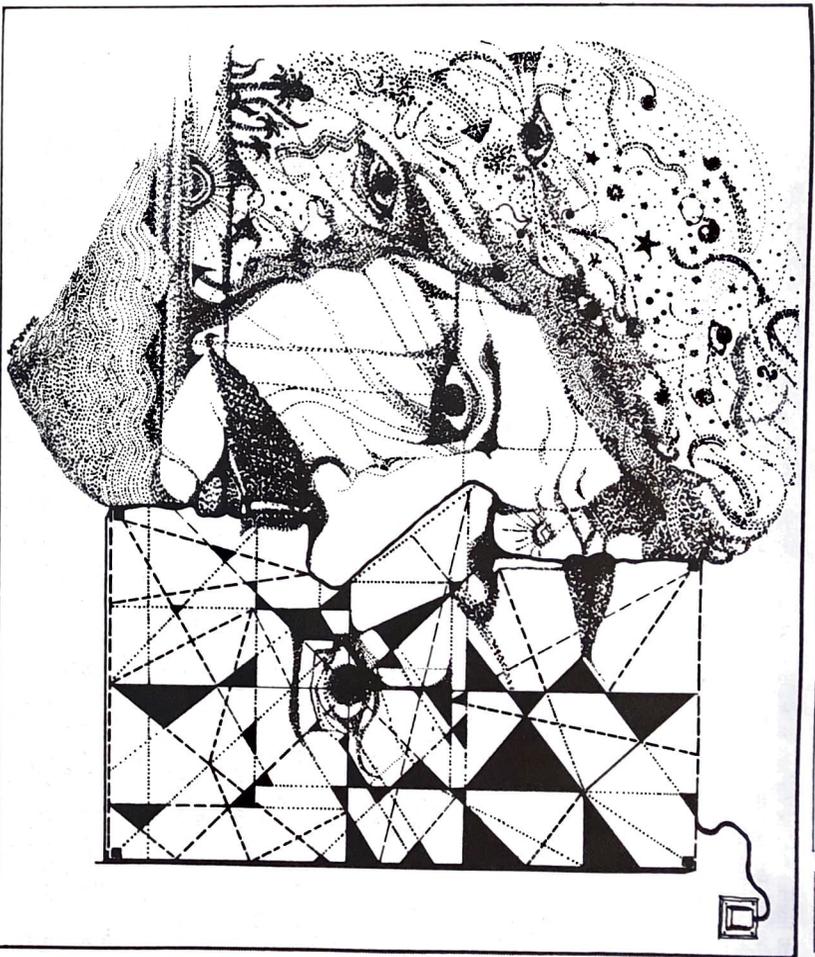
only be realized when we have totally duplicated the brain', p 573, my italics!) Thus, Peter Large, technology editor of *The Guardian*, rightly cautions against our undue reliance on the so-called thinking machines because these rely on two uncertain sciences: 'our knowledge of ourselves and our knowledge of our own creation, the computer'.

One could take our story further, though its main actors no longer are forceful personalities like Wiener and Turing but modest and unassuming collectivities of bland scientists, by saying a few words about the present range and scope of AI activities. Before that, however, some history of the term AI itself is in order. The first published appearance of the term *artificial intelligence* is supposed to have occurred in the proposal made by John McCarthy, of Dartmouth College in Hanover, New Hampshire. His 'two-month, ten-man study', he wrote, 'is to proceed on the basis of the conjecture that every aspect of learning, or any other feature of intelligence, can in principle be so precisely described that a machine can be made to simulate it'. As for its modern manifestations, unquestionably, AI has had its greatest success in the area of game playing. Arthur Samuel's programme of the simple game of checkers, for instance, has been rated as high as that of the order of the top human players in the world. Samuel's program made use of the human strategy of looking ahead and envisioning possible moves in advance. The same approach to the intellectually complex game like chess, however, proved to be quite unsatisfactory. The number of possible moves in a game of chess is too staggering even for a computer to calculate within a reasonable fraction of time. (In a typical chess game the number of possible moves is something like 10 raised to the power of 120, ie one followed by 120 zeros. Supposing that a computer had started working on that task at the moment of the creation of the universe, say some 4.5 billion years ago, then the computer would still be counting and would probably have to work for another life span of this universe to finish its calculations!) Thus alternative strategies like the 'brute force look-ahead' and 'heuristic (rules of thumb) look-ahead' had to be devised. Despite all this, it is far to say that computer chess still lacks the sophistication and ingenuity of its human masters from (?) its wartime birth, just as some of the cultural assumptions in AI technologies have

an unmistakable 'military bias'. (A convincing testimony to this fact is supplied by Blay Whitty of SEAE, Department of Computing and Cybernetics, Brighton Polytechnic, in his paper: 'The Computer as a Cultural Artefact', in Gill, K aramit, S (ed): *Artificial Intelligence for Society*, Chichester and New York, 1986, pp 115-124.)

One area in which AI is increasingly making its presence felt is the so-called sphere of 'expert systems'. Plainly speaking, an expert system is a computer-based application program that exhibits human-like expertise and intelligent behaviour in certain restricted domains such as medical diagnosis, accountancy, legal advice, educational testing and so forth. These programs store 'expertise' as humans and by a skillful question and answer technique are able to, say, successfully diagnose an illness, figure out what is wrong with malfunctioning vehicles, give grades to college students, and perhaps prepare even legal briefs. Clearly, within well-defined fields of professional activities, much can be done to eliminate drudgery and routine from human intellectual labour and transfer it to machines. Nonetheless, even such remarkable success in applied fields hardly elevates AI to the status of a 'science', which is the claim of its most clamorous enthusiasts. Similar is the case in other AI breakthroughs in the fields of vision, such as CAD, interactive Video, or various other experimental activities that involve the identification of printed or written text, even in Chinese and Japanese Characters, or other pictorial identifications.

There do exist, however, certain domains of AI that genuinely impinge upon the question of intelligence and substantiate, however partially, its claim to be a science of mental faculties through the use of computational models. Here may be counted all those activities that have to do with the understanding and production of natural languages (for instance, the once highly acclaimed SHRDLU of Terry Winograd of MIT) and all their ramifications such as translating, making paraphrases of longer pieces of text or even writing abstract poetry like the Japanese haiku. To this list must also be added those experimental domains which try to stimulate human creativity in the arts by computer human creativity in the arts by posing music or prod using pictorial art. Other abstract applications of AI are at present in the fields of analogical thinking, such as proving mathematical theorems, and learning concep-



Graphic artist, Zafar Malik's depiction of the dilemma posed by the advent of Artificial Intelligence. Can there be a meeting point between intuitive logic and cold reasoning?

to man and hence, it can neither be duplicated nor simulated by machines. Then, there was, what may be called rather insensitively as, the *obscurantist* objection because it shied away from arguing its position simply by 'burying its head in the sand' and hoping that such a dreadful thing as a thinking machine would never see the light of day. Beneath this 'obscurantist' sentiment, however, there unmistakably lurks, in my opinion at least, the sound instinctive metaphysical horror of a biological species. Its lack of argumentational sophistication may not be construed as the weakness of its stance. On such a protean issue as that of the mechanical and the human, primordial emotions stirred by biologi-

cal formation and the like. Undoubtedly, computer-based fields of AI activities not only invade every arena of our professional life, they are also growing more abstract and 'human' day by day. Moreover, despite the technological bias of the current research programmes of AI, bequeathed by Turing's functional definition of 'intelligence', there is no reason to believe that its theoretical sub-structure has remained intact since his days. In fact, the truth is that while the idea of stimulating human intelligence was the source of the historical inspiration of AI, it no longer defines the discipline as it exists today. Any philosophical and moral appraisal of AI, therefore, would have to take into consideration, besides its early history and utilitarian present, its latter development and future promises as well. Undoubtedly, the subject of AI elicits powerful emotional, often polarized, response from all of us. At some level of this discourse we sense

a need to define our human identity, either by claiming special attributes which machines do not, indeed cannot, share with us, or by doing just the opposite, namely erasing all ontological barriers between us and machines. In contemplating the prospects of thinking artifices, it would also be true to say, most of us experience some intense form of 'metaphysical anxiety' which makes orderly discussion or calm reflection well nigh impossible. Moreover, despite the novelty of the theme of AI, some of the subjects, indeed arguments, of this debate are as old as human civilisation itself. Indeed, aware of the storm of protests that would arise following his claim that machines could think, Turing himself anticipated some of the likely objections and classified them according to a thematic scheme. At the top of his list, not quite unexpectedly, is the *theological* argument which rests on the axiom that intelligence is the special gift of God

cal instinct may be far less misleading than all the arguments advanced by the acquired intellect and reason of mankind!

Perhaps the most cogent and compelling objection to the mechanistic model of the mind comes from J R Lucas of Oxford. Though Turing had anticipated it in some rudimentary form - and Hofstadter has exhausted all his acumen in trying to refute it - Lucas' argument remains logically and emotionally unassailable - for a certain kind of human personality, at least. Essentially, the argument, which, quite naturally, does not originate with Lucas but is encountered in one pristine form or another in all branches of human civilisation (The poetic genius Jalaluddin Rumi, for instance, brings out this self-evident paradox of self-contemplating thought in his characteristically masterly fashion), rests on the paradox of thought thinking itself, the consciousness of being conscious as it were. Lucas himself expresses it as: 'The paradoxes of consciousness arise because a conscious being can be aware of itself, as well as of other things, and yet cannot really be construed as being divisible into parts. It means that a conscious being can deal with Goedelian questions in a way which machines cannot, because a conscious being can consider itself and its performance and yet not be other than that which did the performance. A machine can be made in a manner of speaking to 'consider' its performance, but it cannot 'take this into account' without thereby becoming a different machine, namely the old machine with a 'new part' added. But it is inherent in our idea of a conscious mind that it can reflect upon itself and criticise its own performance; and no extra part is required to do this: it is already complete and has no 'Achilles heel.' (A more comprehensive statement of the 'materialist' sentiment is found in Herbert Dreyfus', *What Computers Can't Do: A Critique of Artificial Reason*, New York, 1972.)

The annals of mind-machine controversy, however, are virtually inexhaustible and no satisfactory summary can be provided within the scope of this paper. Suffice it to note here is the fact that even if AI no longer makes any theoretical pretence to unravelling the mysteries of the human mind or operate by simulating its activities, such is the emotive force of this subject that it cannot, quite simply, be banished from this discourse. An alternative, and in our opinion more sensibly practical, view to es-

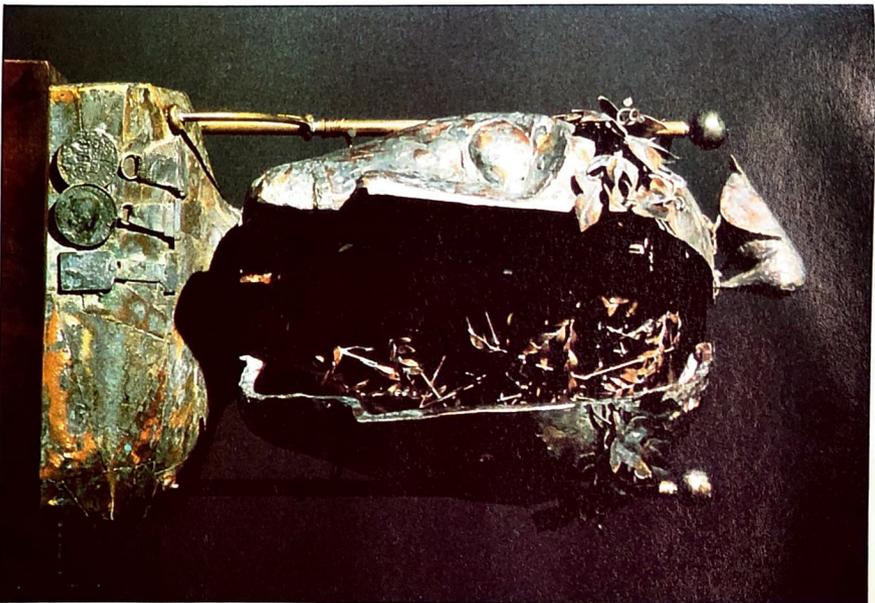
chew, if not terminate, this controversy is provided, in a true Tunngian fashion, by Ajit Narayan of the Department of Computer Science, University of Exeter ('Why AI Cannot be Wrong', in Karamjit S. Gill (ed): *op.cit.* pp 43-54). Narayan's argument, plea actually, is essentially ethical and wishes to bring the multi-disciplinary activities of AI under some form of ethical scrutiny and control. Though born out of the 'epistemological despair' regarding AI's claim to being a science, Narayan in fact makes virtue out of necessity. Disclaiming that AI research satisfies traditional criteria for what it is for a discipline to be a science, Narayan also discards any 'beliefs' that there is some shared principle of knowledge informing various AI sub-disciplines or that it has any 'mind-expanding' epistemological mission. From this critical stance, however, comes the plea that AI be regarded a 'technology' and, consequently, be held accountable for its activities.

With an uncanny intuition about the sacrosanct position of 'science' in modern culture, Narayan argues his case with shrewd and tactful logic: '... if AI is mere technology, any attempt to pursue AI research for the sake of science will not succeed. Much has been made of the dangers of AI to mankind. If AI is a science, AI researchers can ignore these dangers by claiming that science had to be pursued for science's sake, in the same way that scientists working on nuclear fission argued that science had to be pursued. However, if AI is a technology, it can be held accountable and can therefore be constrained (my italics). No doubt, the ethical thrust of Narayan's neo-Tunngian approach to AI would be hailed by all of us. In fact, the distinction posited by him between the ethically and politically accountable 'applied research' and the freely pursued 'theoretical science' could have practical consequences far beyond the bounds of his own discipline of AI research. Nonetheless, it is also fair to point out that such a contra-distinction can only function upto a point and is forever subject to political and academic manoeuvring, simply because it does not present any rigorously theoretical internal criteria for the differentiation of 'science' from 'technology'. Theoretical and conceptual uncertainties, it must be emphasized, cannot always be bypassed by practical strategies. And that goes for AI as well. Its moral issues must be faced both intellectually and 'pragmatically'.

The interface of ethics and AI, even if we disregard the highly emotive issue of mind vs machines, is of such critical significance that not only does it demand much careful reflection and thought from the adherents of every school of human thought - Muslims including - but it also raises such disconcertingly fundamental issues that it is inconceivable as to how it may be deliberated without asking seminal questions about our own human identity and destiny - problems that are regarded as 'meaningless and insoluble' in all modern discourse! AI paradigm, moreover, may be understood in two different ways: the narrow claim according to which only those segments of human mentality that may be defined quite stringently as 'intelligence' can be reproduced with total fidelity in computer programs, whilst the wider claim has it that all aspects of mentality can in principle be realized on computer systems of some desirable degree of complexity. The difficulties of this 'ambivalence' have been very aptly described by Steve Torrence of the University of Sussex (*ibid.*: Gill, K.S.: *op.cit.*, p 57) as such:

'It is thus certainly very easy to slip from a narrow, relatively focussed, sense of 'intelligence' to a wider, vaguer notion, a notion which seems to encompass the whole of mentality, including desire, emotion, direct conscious experience, pleasure, pain, etc. Because of this ready transition, the field of artificial intelligence seems to take on a much more portentous air. AI theorists become taken in by their own sleight of hand. We start out with a very limited and plausible enough claim concerning a particular set of cognitive activities - namely that computer performances of such activities may be called intelligent in exactly the same sense in which the human performance of those activities are. We end up with a grand theory of mentality as such - that any mental state, process, activity, capacity whatsoever is in principle computer simulable, and therefore computationally explicable'.

Seen in this light, then, the issue does not present itself as a mind-machine alternative but a mind-machine continuum. The question that arises then is: In which sense may AI be made an adjunct to human brain and how may the two come together in a new synthesis? For some, the very posing of the question is enough to make them lyrical. Here, for instance, is one answer: 'If we can fuse AI with the human mind, then those two



areas that conform to the limited data they are capable of programming. AI and other computer software have indeed ushered us in an era where creativity has become synonymous with computer compatibility. What cannot be programmed is simply not respectable enough in the world of business, arts, learning, and culture! The writer who submits his manuscript type-written rather than laser-printed, the architect who presents his designs without the aid of CAD, the entrepreneur who exhibits his projections minus spreadsheet graphics and so forth, all run the risk of being dubbed as amateurs. AI, with its very limited spectrum of programmable human mental activities, promises to aggravate this development even further. Intelligence already has come to signify computer intelligence and the Turing test of being intelligent is to be able to write 'intelligent' programmes! No matter what the defenders of AI may say about its humble claims, the fact is that AI is seriously constraining human creativity by default if not by purpose.

The Brave New World of AI is ruled by the demi-god of automation. Man need not do what may be accomplished by the machines, seems to be inscribed in its holy creed. In this narcissistic universe of self-gratification, there is no room for self-sacrifice. Little wonder that the instrumentalist approach to thinking and the computational approach to intelligence have today triumphed over the human-centred ideals of self-transcendence. Today, the technical mind is making its impudent assault on the last bastion of human autonomy - the human mind. If man earlier could take comfort in his being free in mind and thought, he has reasons to be apprehensive today. What AI aims at is not the simulation but the replacement of human thought. One of the most somber insights of our times, expressed by Lewis Mumford, sees in the perfection of automation the death of man. Without a trace of pretence and inhibition he makes plain his claim that, 'Once complete and universal, total automation means total renunciation of life and eventually total extinction: the very retreat into Nirvana.'

All the pretence and ploy of Artificial Intelligence, canonized as it is at the cost of denying man his unique gift of intelligence, is undoubtedly the Nirvana of the technical man. Will he come out of it as the Enlightened One or will he simply become extinct, that is the question. ■

great bodies of knowledge will come together in a single man-machine system - a binary brain. The computer will provide virtually endless and infallible memory, plus prodigious power of data crunching. The human brain and mind will supply all that special know-how that a human acquires... Each part of the system will give the other something it desperately needs but lacks. And the two avenues of evolution - Lamarckian and Darwinian - which diverged with the invention of the first computers, will come together again, like the themes of a fugue at the close. The results will be more awe-inspiring than we can imagine. We can no more envision the deepest workings of such an intelligence, than a dog can understand the stars.

Fortunately, most of the saner voices of humanity have cautioned against just such a kind of infantile technological optimism. The prospects of mind-machine marriages have generally not elicited much enthusiasm even among the high priests of technology. Wiener himself, for instance, revered man's autonomy, his unpredictability and moral responsibility inestimably. No one better understood the problems that the autonomous intelligence systems independent of human intervention would create, and no one was more apprehensive than he over the frenidish fascination automated systems would have for automatic minds. He was fully aware of the danger that petty programmers with gigantic egos would be eager to confine human reactions to those