

technology has reified science fiction and brought fantasy into our everyday lives. The computer is a tool of the mind allowing us to explore our wildest imaginations; it is also a tool too close to ourselves for comfort, perhaps replacing us in activities that we have been taught are our purpose for existence, such as work.

How can we encapsulate the complexities of computing technology and see it in terms that we already know? One view that has already proved very useful is to regard the computer as providing us with a new communications medium supporting interaction and conversation. Books are remarkable in allowing voices from the past to communicate with us and influence our emotions, values, knowledge and actions. But books are one-way. An author may come alive to us, as may his characters and his concepts, but we cannot answer back or ask questions.

Computers provide a medium analogous to a book that supports two-way communication. The programmer may simulate himself, the personalities of others, and the phenomena of complete worlds of objects and of people that may, or may not, have ever existed. Adventure games on the personal computer in the home allow us to become a character in an electronic "novel". We learn to live in an alien environment, gain friends and resources, and use them to achieve our goals. Expert systems on larger computers allow the recorded mind of an expert, or the composite "mind" of many, to guide others less skilled in complex tasks. In using them you are discussing with a "colleague" problems of medical diagnosis or oil exploration, or the inventive processes of mathematical discovery. You tell them your problem and the information you have. You discuss it with them, query their judgement and ultimately come to a decision based on a collaboration with someone who may be long dead.

Computer technology affects the core of our being because conversation is at the heart of human civilization. We are a species remarkable for our adaptive and learning capabilities, the effect of which is immensely amplified by our capability for conversation. Only one person need learn from experience. Others can be told and media allow the telling to transcend space and time. New media *do* change our world and the computer provides the first one for encoding not just a conversation but the capability to converse itself, not just a picture but the capability to be in the world portrayed. Pask and Curran have made superb use of an existing medium to give an exquisite portrayal of that which is to come. □

Brian Gaines is Professor in the Department of Industrial Engineering at the University of Toronto, and Mildred Shaw Associate Professor in the Department of Computer Science at York University, Ontario, Canada.

Futuristic futures

Margaret A. Boden

The Super-intelligent Machine: An Electronic Odyssey.

By Adrian Berry.

Cape: 1983. Pp.180. £7.95.

The Intimate Machine: Close Encounters with the New Computers.

By Neil Frude.

Century, 76 Old Compton St, London: 1982. Pp.190. Hbk £8.95; pbk £4.95.

THE Duchess assured Alice that "Every thing's got a moral, if only you can find it". The moral of these two books is not hard to find. They both offer a taste (a soupçon) of what has already been achieved by artificial intelligence, but their main message in each case is: "You ain't seen nothing yet". The specific "nothings" the authors have in mind, however, differ.

Berry caps a highly superficial account of some current programs with the forecast that super-intelligent machines will inevitably surpass and replace us. They will outlast all conceivable biological species, which are limited by thermodynamics: long after the Universe is too cold for life-forms to persist or to be formed, intelligent computerized systems could continue to evolve. They could even forestall their own annihilation, by devising a "technological fix" on a literally cosmic scale to prevent the otherwise inevitable collapse of the galaxies into one infinitely massive point.

Berry admits that this is stranger than the strangest science fiction, but his readers' respect for Wells and Asimov may not suffice to settle their doubts. What I found no less difficult to accept is the vulgarity of style and superficiality of argument throughout the book. The question "Will time end, or will it go on for ever?" is given short shrift: the Big Bang theory is cited as showing definitively that time had a beginning, and then our question is answered for us, since "What had a beginning may surely have an end". Again, instead of discussing Abelson's pioneering "Goldwater Machine" (to which he merely refers), Berry tells us that he himself wrote a "Wedgwood Benn Machine", whose quoted conversation provides an easy crack at his *bete noire*, but which casts no light on the way in which computer models of belief systems actually function. An author who is science editor of the *Daily Telegraph* could surely have served his readers better.

Frude's book, which is the more original and interesting, extrapolates from those current programs which embody some sort of "friendliness" and/or "personal" characteristics. These include simulations of political or paranoid belief systems, systems capable of automatic speech-recognition and speech-production (the latter with increasingly realistic vocal timbre and intonation), and computer models

of patterns of social interaction and of distinct personality-types.

The author is a social psychologist, interested in the forms of interaction that may develop between human beings and computers. He argues that — irrespective of whether a programmed system could in principle be *really* friendly — appropriately-programmed computers will increasingly *appear* to be personal beings, and will be so taken by their human users, who will experience interaction with them as a form of intimate relationship. He reminds us that even today's crude beginnings are viewed personally by many people, and that users often adopt an inappropriately anthropomorphic stance towards computer systems. His main point is that this is due not merely to a lack of computer literacy, but to a universal human tendency toward animism, and that this response is more likely to wax than to wane. With increasingly human-like performance on the part of programmed systems, our natural animistic tendencies will often swamp our scientific understanding (which for many users may be minimal in any case).

It is a rare adult who has never smiled back at a smiling teddy-bear, or winced when it was punched on the nose by an unfeeling (*sic*) person. But these sorts of dolls are mere child's play. The novels of Tom Sharpe, if not our own experience, have alerted us to the existence of less innocent sorts of doll: rubber inflatables sold by sex-shops for unprintable purposes. Frude sees it as a commercial inevitability that, among the many friendly and convivial programs to be developed in the future, some will be embodied in "soft machines", sexually seductive or merely cuddly depending on the market concerned. Young monkeys spend more time clinging to a non-feeding monkey-doll that is covered with soft towelling than to the wire-frame "mother" which provides them with milk. Our comparable preference for softness and warmth, combined with our animism, could lead us to have significant emotional engagements with the dolls of the future — whether suited to sex-shops or not.

Past achievements in puppetry, impressive as they are, will be as nothing to the programmed versions. Already in 1970, a soft robot was exhibited in Japan that had 30 distinct "facial muscles" to generate expressions. A wide range of emotions were attributed to this creature by its human audience. Slight changes of speed or muscle-combinations often led to significant

Fresh yeast

Cold Spring Harbor Laboratory have published the companion volume to *The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance* (reviewed in *Nature* 299, 92; 1982). The second volume, Monograph 11B, is subtitled *Metabolism and Gene Expression* and costs \$75 in the United States, \$90 elsewhere.

changes in the interpretations given to the robot's behaviour; for instance, a pleasant smile executed at half-speed appeared as ugly and menacing. Obviously, it would have been a relatively simple matter to provide the robot with artificial tear-glands, the tear-ducts being opened only when the appropriate facial expression was being displayed. Frude points out that future tearful statues of the Virgin could appear even more miraculous than any past attempts at religious trumpery.

"Trumpery" here is my term, not Frude's. In general, he is curiously coy about the moral evaluation of these predicted developments, or at least sees this as more ambiguous than many others would. He forecasts the end of social isolation, when people can buy "a friend off the shelf", and the alleviation of domestic suffering given the widespread use of programmed marriage-counsellors and anti-suicidal Samaritans. Religious and political persuasion and counselling, too, will be available, untouched by human hand. He does not seem to be speaking with tongue in cheek: there is hardly a suggestion that the "social" communication and "religious" counselling he mentions are different in principle from those we enjoy today. There will of course be some deficiencies in practice, due to insufficient advances in the technology; but these, he says, will be largely compensated for by the human users — much as we compensate for the abnormal speech of a foreigner, or of a laryngectomy patient. And he repeatedly rebuts the general question of whether intimate computers *should* be written, always with the claim that irresistible commercial pressures will make them widely available, whether we like it or not.

The general reader may be misled into believing that an impressive artificial intelligence is only just around the corner. Some forward-looking businessman, dazzled by Frude's forecasts of vast profits, may even rush to beg a loan from his nearest merchant-banker — though, bankers being cautious characters, his mission should be fruitless. Neither author sufficiently emphasizes the difficulties involved, the enormous size of the gap between current programs and the human mind. Certainly, Frude does alert his readers to these difficulties every so often. After describing the smiling robot, for example, he points out that "Getting a machine to laugh is easy. Getting it to laugh at a joke is very, very, difficult". Berry, by contrast, writes as though the difficulties were negligible, as though our evolutionary replacement by the first non-biological species — a species capable of influencing the nature of the end of time itself — is a clear certainty. But then, seeing a joke may be difficult for human beings, too. □

Margaret A. Boden is Professor of Philosophy and Psychology at the University of Sussex. Her books include Artificial Intelligence and Natural Man (Basic Books, 1977).

Thoughts of sorts

Stuart Sutherland

Science and Moral Priority: Merging Mind, Brain, and Human Values.

By Roger Sperry.

Basil Blackwell/Columbia University Press: 1983. Pp.135. £12.50, \$16.95.

MANY years ago I worked briefly in Professor Roger Sperry's laboratory at Caltech. I was running experiments on goldfish and on cats, and from time to time he would pop his head round the door, enquire "Have you stopped feeding the goldfish to the cats yet?" and disappear with a chuckle. The chuckle quotient of *Science and Moral Priority* is low, and it is apparent that having won a Nobel Prize for his elegant research on the development of the nervous system and on the split-brain, Sperry has turned his mind to higher things, namely, the nature of consciousness and the selection of moral values. From Charles Sherrington to John Eccles it has been characteristic of neuroscientists to devote their declining years to metaphysical problems. It is also characteristic that they have refused to burden themselves with a knowledge of any previous work on the subject, thus leaving themselves all the freer to speculate. Sperry's new book upholds this tradition.

On the mind-body problem, he adopts a monist position, in which consciousness is an emergent property of the brain. For him, consciousness is made up of complex patterns of activity in the nervous system which apparently can be understood only in mental not physical terms. He believes these patterns of activity are identical to perception, thinking, taking decisions and so on. Since they determine other events in the nervous system and control the muscles, he ascribes to the mind a causal force. He appears to be torn between his scientific view of the brain and his desire to preserve the importance of consciousness, since on page 68 he writes that consciousness is not reducible to neural events whereas on page 96 he states that it is.

Sperry's main and frequently repeated argument for his form of monism, which he modestly describes as a "conceptual breakthrough", is that it constitutes a "paradigm change" that is widely accepted by the informed majority. Leaving aside the danger of deciding truth on a ballot, it is not true that the majority of cognitive and brain scientists, whether informed or not, adopt this view of the nature of consciousness. Most of them are wise enough to hold no view and the fact that the phenomena of consciousness have received increased attention over the past 30 years has nothing to do with anyone's opinions of their status. Indeed, Sperry does not attempt to specify how this "paradigm change" has actually influenced scientific research on cognition and the brain, nor could he do so

since, provided consciousness is in some way linked to brain activity, its status is irrelevant. He is right in thinking that any higher-level system can be explained only in terms of concepts appropriate to that level: the behaviour of molecules cannot be explained solely in terms of the concepts that apply to individual atoms nor can the computations carried out by a program be explained solely in terms of transistors. But this merely reflects a limitation of the human mind, which cannot manipulate a large number of separate entities at once and which must therefore construct higher-level concepts in order to apprehend higher-level systems that arise from the interactions of more basic entities. Moreover, Sperry's conclusion that the workings of the brain can only be explained in terms of conscious processes is incorrect. It is at present wholly unclear whether the everyday concepts used to describe consciousness will turn out to be the most appropriate ones to apply to higher levels of organization in the nervous system.

Despite his defiance of William James's sensible injunction that "Dualism is a fundamental datum. Let no man join together what God has put asunder", Sperry is honest enough to admit that his thesis has no empirical support, though he resists the implication that it has no consequences empirical or otherwise. Unfortunately, the only consequences he specifies turn out to be vague promises. According to him, his view of consciousness will "logically dissolve . . . the classic fact-value and naturalistic fallacies of philosophy" (whatever they are). Sperry believes that his identity theory of consciousness and brain states (which incidentally is hardly new) will not only make philosophers redundant, it will allow brain scientists to take their place, for the influence of his theory has apparently been such that neuroscience is no longer "dehumanizing": hence the way is open to found human values "on a reference framework based on empirical evidence and the scientific method". He gives two main arguments for this curious view.

First, he believes that since science has proved the most successful method for discovering empirical truth, it must also be the best way to select moral ends. Even the premise of this non-sequitur may be doubted, since there are many kinds of truth and it could be argued that literature has implicitly revealed much more important truths about human values than has the scientific study of man. Moreover, Sperry does not disclose how science can reveal moral ends. He adopts, without acknowledgement, a rather broader version of C.H. Waddington's view that it is morally right to promote the ends (whatever they are) of evolution. Sperry talks about the "upward thrust" of evolution, without realizing that the word "upward" implies a value judgement that cannot be derived merely from the scientific study of evolution. He seems to include in evolution