

## The chips are now down?

Steve Woolgar

**Mind Over Machine: The Power of Human Intuition and Expertise in the Era of the Computer.** By Hubert L. Dreyfus and Stuart E. Dreyfus. Basil Blackwell/Free Press: 1986. Pp.231. £15, \$16.85.

THE most important and exciting thing about artificial intelligence (AI) is that it provides the possibility of a critical test of fundamental assumptions about the nature of knowledge and reason. These assumptions have gripped Western thought for 2,000 years and they underpin ingrained ideas about the character of man, his abilities and potential. If the project to devise artificial reason is adjudged a success, it will vindicate the objectivist and rationalist philosophies which are thought to have guided the rise of science and technology; its failure will buttress those phenomenological and interpretivist philosophies which argue that calculation and measurement are largely irrelevant to man's capacity for intelligent behaviour.

Hubert Dreyfus recognizes what is at stake. His *What Computers Can't Do*, first published in 1972 and revised in 1979, was a compellingly sophisticated analysis of the philosophical assumptions which underlie AI. That book provided an overview of developments in AI from work on cognitive simulation in the 1950s to early attempts at knowledge engineering in the 1970s. *Mind Over Machine*, written with his brother (a computer specialist), extends the analysis to recent work, particularly on expert systems. The book's appearance is timely, in view of the sudden and massive increase in financial support for AI, especially from military sources.

Will these people collect on their expensive bet? The Dreyfuses say no. Their central thesis is that intelligent human behaviour (interpretation, understanding, recognition, puzzle solving, rule application and so on) is simply not amenable to formalization in the way required by digital computers.

The authors propose a model comprising five stages of skill acquisition, ranging from novice to expert. Progression beyond the third stage ("competence") requires a qualitatively different approach to reasoning. Expertise, in their view, involves the accumulation of experience, the unconscious holistic recognition of the familiarity of new situations and the critical use of intuition and "know-how". This is beyond the ability of systems designed on principles of calculation from discrete facts and codified rules. The apparent

success enjoyed by some artificial devices occurs in restricted (artificial) areas which require competence rather than intelligence. This is not to deny the usefulness of some AI systems as aids, but such systems are properly termed "competent" rather than "expert". The five-stage skill hierarchy is used to discern the limits of computer use both in education and in managerial decision-making. Only in the rare instances, where researchers abandon the mechanistic models of human skill, do we find hope for a better approximation to human expertise.

The anti-mechanistic view of human skill derives from the Wittgensteinian argument that the determination of interpretation, understanding, recognition and so on is entirely construed in a social context and can never sensibly occur in isolation from it. But the Dreyfuses' hierarchy of skills is strikingly individualistic and mentalistic. For example, they characterize expertise in terms of inner processes: "the expert driver becomes one with his car, and he experiences himself simply as driving, rather than as driving a car . . ." (p.30). Given the philosophical orientation of their critique, there is surprisingly little acknowledgement of the way in which such experiences are meaningful only as socially sanctioned and publicly accountable descriptions. Although the authors insist that their aim is not to mystify, their treatment of skills as inner processes and experiences provides too easy a target for the rationalists. Similarly, it is surely offering too much of a hostage to fortune to continue to speak of the brain and the mind as receptacles of intelligence. The Dreyfuses intend these curious and outdated concepts as a sanctuary for unformalizable human skills. But as others such as Jeff Coulter have shown, notions of "brain" and "mind" are social constructions which tend to support the mechanistic view of human behaviour.

The Dreyfuses are especially impressive in demonstrating that despite changes in emphasis and direction (for example, the shift from attempts to fashion general,

problem-solving devices to those able to make inferences within a highly specialized knowledge domain), workers in AI repeatedly face the same fundamental conceptual problems. For the Dreyfuses, AI's failures are the direct result of such irresolvable problems. For many AI enthusiasts, however, these are merely temporary, technical and surmountable difficulties. For example, whereas the Dreyfuses say that work on "expert" systems shows it is wrong to assume that human experts are using heuristic rules, proponents of AI interpret the research as merely indicating some snags in eliciting these rules. (It is notable that most previous responses to Hubert Dreyfus's ideas have preferred to contest the alleged failure of particular systems, rather than address the conceptual problems he raises.)

Early restricted successes gave rise to unwarranted optimism about the future of AI and, for the Dreyfuses, this optimism partly accounts for the current inability of the AI community to recognize failure. But the Dreyfuses also suggest that the stakes are now so high that failure is actively concealed. They thus allege deliberate cover-up by the AI establishment, describing in the book their own treatment and the manipulation of TV appearances by the "artificial intelligentsia", and criticizing the "unscientific" attitude of AI workers. (The dust-jacket speaks of "AI's record of barefaced public deception".)

This recourse to a kind of conspiracy theory is something of a distraction. For a full appreciation of the dynamics of the AI enterprise, we need to look carefully at the social attribution of "success" and "failure" in the context of a pervasive commitment to rationalistic models of intelligence. Although the Dreyfuses' conceptual argument provides a good, sceptical springboard from which to launch a social analysis of AI research, we should avoid their supposition that conceptual problems will necessarily give rise to "failures".

Who, in any case, are those who cover up failures and spread false optimism about AI? It is important to be clear that the Dreyfuses identify a relatively small body of researchers. True, these are the most influential and respected champions of AI and are thus a legitimate and necessary target for attack; many are entrepreneurs with a substantial interest in furthering public support for the enterprise. But it might be misleading to accept their proclamations as characteristic of AI as a whole. Searle's distinction between "strong" and "weak" AI is useful here. Whereas proponents of strong AI (who claim, for example, that programs can "understand") are appropriate targets for the Dreyfuses, the more cautious advocates of weak AI (who see computers as,

IMAGE  
UNAVAILABLE  
FOR  
COPYRIGHT  
REASONS

at best, a tool for aiding our investigations into the nature of intelligence) might agree that it is unreasonable to expect the success of artificial systems beyond the level of competence.

Weak and strong AI are not unrelated, of course. Proponents of the former have undoubtedly benefited from the public persuasiveness of the latter. But the distinction points to a further shortcoming in the Dreyfuses' case. By concentrating upon proponents of strong AI, the arguments give us little insight into what practitioners of AI do. In this book, the Dreyfuses continue their impressive assault in the titanic struggle between phenomenological philosophy and the rationalist assumptions of AI's public guardians. But we are not told whether or how the assumptions of strong AI touch upon the day-to-day business of devising systems, making programs work, evaluating output and so on.

It is interesting to consider the likely fate of the Dreyfuses' thesis. Their case against AI is the strongest and most eloquent yet published, a protest against conceptions of the human being which have held sway since the beginnings of Western thought, and it suggests that one outcome of the AI debate might be a definitive refutation of the rationalist metaphysics. However, the lack of a fully social conception of the practice of machine intelligence is perhaps reflected in their inattention to one of the key lessons of the history of science: "critical tests" are rarely decisive. Evaluations of "success" and "failure" need only generate further redefinitions of AI, not its abandonment. □

*Steve Woolgar is a Lecturer in the Department of Human Sciences, Brunel, The University of West London, Uxbridge, Middlesex UB8 3PH, UK. With Bruno Latour he is author of Laboratory Life: The Construction of Scientific Facts, a revised edition of which has recently been published by Princeton University Press.*

## Wondrous gifts

Peter Bryant

**Nature's Gambit: Child Prodigies and the Development of Human Potential.** By David Henry Feldman with Lynn T. Goldsmith. *Basic Books:1986. Pp.259. \$19.95.*

EVERY now and then a child turns out to be outstandingly talented in a quite specific way. It is not just that he quickly outstrips all his contemporaries in a particular skill: he also soon becomes a great deal better than most grown-ups at it. Children such as this arouse great interest and wonder. Psychologists in particular have to pay attention to them because at first sight they seem to break most of the rules set out by developmental theories. After all, these theories generally hold that different intellectual skills are connected and that they develop together at roughly the same pace through childhood. How can such theories possibly explain a five-year-old child whose mathematical skills already outstrip those of 95 per cent of the adult population, but who otherwise seems pretty ordinary?

I suspect that this question was the starting point for David Feldman's book, which is a prolonged and detailed study of six different, prodigiously talented American boys and their families, but it is not in the end a question that he answers satisfactorily. Instead he spends most time on another question which is unanswerable, or at any rate cannot be answered from a study of individual child prodigies. His main concern is to find what conditions are needed to produce such children. Feldman's own idea, which is about what he calls the "coincidence process" is neither new nor sophisticated: it is that these

children have extraordinary gifts ("proclivities" is Feldman's term) to start with but that their talents only flourish in the right environment. Prodigies need sensitive and devoted parents and an appropriate social environment, according to Feldman, and he is plainly impressed by the lengths to which some of the parents in his little group were prepared to go to support their children. Sacrificing jobs, commuting long distances, confronting awkward teachers, taking the child from one tutor to the next — these seem to be typical demands for a prodigy's parents, and in this group at least they shouldered the burden quite cheerfully. Often the parents' own talents matched their child's: the boy with an extraordinary gift for music had a musician for a father, and the father of the boy who showed early signs of being a talented writer was a writer himself. In these cases the parents spent a great deal of time teaching their own specific skill to their child.

Feldman's argument that all the children in his group were lucky enough to be born into a suitable environment is convincing, but it does not prove his case that such environments are essential for the emergence of a prodigy. The trouble is that his thesis is just as much about children whom he did not study as about the six prodigies with whom he spent so much time. There must, he argues, be children who could have been prodigies, who had the "proclivities", but who found themselves in the wrong environment: the world must be littered with mute, inglorious Miltons or Cromwells guiltless of their country's blood. But Feldman did not study any such prodigy *manqué*. He does not even demonstrate that they exist, and it is very difficult to see how he could have done so, given his methods.

IMAGE  
UNAVAILABLE  
FOR  
COPYRIGHT  
REASONS

*Making music — the prodigiously talented Mozart, accompanied by his father and sister.*

Feldman's description of the lives of the six boys is interesting and frequently poignant: but it is difficult to see what is new here. After all, we already know a great deal about the childhood of erstwhile prodigies. Mozart's childhood, to take one example, is documented as well as any other part of his life. Sartre, to take another, wrote about his early years in great detail and with considerable distaste. One of the most striking features of these examples and the children studied by Feldman is the surprise and interest that the prodigies excited in the adults around them. The children soon become performers and they do so at the insistence of adults. On occasion, adult astonishment at the child's unexpected talents has some odd results. In my view the most interesting part of Feldman's book is his description of the way in which parents or teachers sometimes resort to irrational explanations for a child's extraordinary talents, because they cannot account for them in any other way. Reincarnation is one solution. The child is thought to be so talented because he is simply a new version of a dead genius.

To read Feldman's sympathetic and sensitive accounts of these six children is to feel the same kind of awe which the young Mozart must have inspired in those who heard him play and improvise so effortlessly. Readers of this book will wonder at the phenomena which Feldman describes, but will remain puzzled about the reasons for them and about the implications for developmental psychology. □

*Peter Bryant is Watts Professor in the Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, UK.*