ARTIFICIAL INTELLIGENCE AND THE BIOLOGICAL FACTOR

Andy Clark

1985

Cognitive Studies Research Papers

Serial No. CSRP. 049

The University of Sussex,
Cognitive Studies Programme,
School of Social Sciences,
Falmer, Brighton BN19QN

SSX
041

ARTIFICIAL INTELLIGENCE AND THE BIOLOGICAL METAPHOR.

ANDY CLARK,
COGNITIVE STUDIES RESEARCH PROGRAMME,
SCHOOL OF SOCIAL SCIENCES,
UNIVERSITY OF SUSSEX,
BRIGHTON, BN1 9QN
ENGLAND.

1985

ARTIFICIAL INTELLIGENCE AND THE BIOLOGICAL METAPHOR.

Introduction.

The computational metaphor in psychology says that it is often useful to consider the mind as a computational system. Guided by this metaphor, much current work in Artificial Intelligence (AI) aims at improving our understanding of human psychology by devising computational models of human competence. In what follows I suggest that such work may usefully benefit from attending to the natural origins and environmental embedding of the mind it seeks to model. I thus urge a biological metaphor as a guide to the proper use of the computational one. In particular, I look at three ways in which biological reflections may influence the choice of problems and the kinds of solution in psychologically-motivated AI. They involve;

1./ The choice of 'micro-worlds' for study;

2./ The kinds of processing strategy we employ;

3./ The exploitation of ambient environmental information.

First, though, I shall say a little about the kinds of worry the adoption of the biological metaphor is meant to resolve.

Cognitive Wheels.

Daniel Dennett has coined the term 'cognitive wheel' to refer to

> "Any design proposal in cognitive theory ... that is profoundly unbiological, however wizardly and elegant it is as a bit of technology".
>
> Dennett (1984) p. 147

The wheel image is meant to capture the fear that current strategies in AI may yield models of mind which bear as little relation to the way natural intelligence works as the wheel bears to natural ways of getting around. Dennett is not alone in these fears. They are shared by workers in the field such as David Marr and Patrick Hayes. Thus we find Marr worrying that AI may 'degenerate into the writing of programs that do no more than mimic, in an unenlightening way, some small aspect of human performance'. (Marr (1977) p. 139. A similar fear is voiced in Hayes (1979) p. 244.

The problem, in effect, is the inevitable tendency of AI to treat the mind as a <u>black box</u>; a system which is known to govern a certain output given a certain input, but whose internal workings remain a mystery. It seems, however, that we do know something which may constrain the kinds of internal workings which we may reasonably postulate. For we do know that the mind is a <u>naturally occurring</u> black box. The biological metaphor urges us to respect that knowledge in the formulation of our computational models of mental activity.There are many ways in which attention to the natural roots of mind may usefully guide and constrain work in AI. I shall look at three such ways.

1. <u>The choice of 'micro-worlds'</u>.

Psychologically motivated work in AI is often defended by means of an optimality argument. Thus suppose we seek to understand how the human brain sorts a series of numbers into order. One strategy would be to find the simplest, fastest sorting algorithm we can and then to claim that human beings probably use that algorithm since our brains, being products of stiff evolutionary pressure to get fast, accurate results, will tend towards the optimal solution to any given problem. But this, I contend, would be to use the optimality argument in a biologically indefensible way. There is, indeed, a kind of modest optimality assumption which may be applied to a certain type of cognitive evolutionary product. But we must be very clear about the domain in which the appeal to that assumption is legitimate. It is legitimate to appeal to optimality only when we are considering very low-level procedures which are of great importance to the survival and success of the creatures which use them. For it is only in this context that we can imagine a small increase in speed and/or accuracy (the two may need to be carefully balanced - see 2 below), conferring an increase in fitness and hence spreading throughout a population. Evolution moreover, is a stage-wise and accumulative process. Old solutions will be adapted and modified to cope with new problems. So often, where complex problems are involved, the natural solution will be very far removed from a computationally optimal one. Rather, it will be a hastily cobbled-up mixture of old strategies pressed into service in a new domaih. In short, the most we can reasonably expect is that evolved creatures will have achieved elegant, perhaps even optimal, solutions to low-level but important problems (such as spotting food and predators etc.), and will combine these solutions in increasingly clumsy and non-optimal ways to solve higher but less survival-relevant problems (such as sorting numbers into order). It would, however, be incorrect to say that the low-level problems are the simple ones - 'simple' seeing involves millions of computations. But they are problems which had to be solved early on in the evolution of natural intelligence and it is plausible to expect our solutions to subsequent problems to build on the strategies developed to deal with these basic requirements.

These observations have an important bearing on the choice of task domains for psychologically motivated AI. It is often sensible to restrict the reasoning which a program aims to model to a narrow, and hence manageable, domain. Thus Uinograd's (1972) SHRDLU is jCLanceiJ3ed^)Dl>LJwith. tbe^siackiJig^QJL^ocks in a 2D blocks world... This 'microworlds' [1] strategy is surely a sound one; we cannot expect to solve all our problems at once. But the choice of a specific high-level feat, such as the stacking of blocks or, in a more recent example, the answering of questions concerning stories about restaurants (Schank and Abelson 1977), may be a mistake. For if the aim is to model human intelligence, then we must bear in mind that intelligence has evolved as a means of satisfying our basic survival requirements. It has <u>not</u> been selected for its capacity to achieve the high-level mental feats which so much work in AI is dedicated to modelling. If we <u>can</u> perform such feats as reasoning about blocks and restaurants, it is only <u>in virtue</u> of our being endowed with a set of low-level capacities which just happen to facilitate the higher-level activity. The AI theorist who goes straight into the attempt to model the high-level achievement is asking for trouble of an all too familiar sort. For if we aim to design a system precisely to achieve a high-level goal, it seems quite likely that some streamlined, isolated strategy will do the trick - a strategy which, however, may bear no relation to any strategies aimed at achieving the basic goals which drive the process of natural selection by which the human mind was fashioned.

The microworlds strategy, applied to high-level cognitive feats, is thus almost certain to yield nothing but cognitive wheels. One AI theorist who sees this fact is David Marr. He suggest that it would be better to focus on low-level intellectual functions and cites work on the visual orientation of the house-fly as an example (Marr (1977) p. 132-3). Attention to the biological metaphor can only support this conclusion. The right microworlds to study are not fragments of the sophisticated human achievement, but the less sophisticated achievements of the various animal intelligences, ranged across the phylogenetic tree. For human intelligence is best seen as a product of the evolutionary accumulation and ad hoc combination of elegant solutions to the simple but important problems which faced our non-human ancestors. In the microworlds of animal intelligence,then, we may expect to find the natural joints on which to carve our computational investigations.

2. Processing strategies.

My colleague at Sussex, Aaron Sloman, has pointed out the way in which the natural need to make time-critical decisions may affect our conception of the global architecture of an intelligent system. He points out (Sloman, 1985) that the natural environment cannot buffer important information until we are ready to receive it. Such information may be vital to our continued survival(e.g. there is a lion about to leap on you). A natural intelligence, then, will need to run various large-scale processes in parallel. It will need to attend to the job in hand all the while being on the lookout for new data which may have to generate an interrupt signal forcing us to drop what we are doing and act on that data. This action will often need to be very fast. Heuristics for quick, fallible decision-making will thus often be preferred to slow accurate algorithms. Natural intelligence will thus utilise strategies which generate decisions very quickly and often on very incomplete informational bases. In a wide range of circumstances, then, no optimality assumption can be justified. The computational modelling of natural intelligence may thus require such non-standard components as fast,fallible heuristics and large-scale parallel processes capable of interrupting one another with new information when appropriate.

3. Exploiting environmental information.

Evolutionary pressures will favour cognitive strategies which are cost-efficient and flexible. One result of this is that the biological metaphor undermines any non-essential use of richly constructivist strategies in AI. By a constructivist strategy I mean a strategy which adds to the incoming information in order to form a rich representation of the world outside the organism. Certainly, the biological metaphor allows for the possibility of our having evolved just such constructivist strategies. This is stressed in Lorenz' (1941) reformulation of the Kantian categories as evolved species a posteriori mechanisms which structure and interpret the base data received by an individual. It is less commonly noticed,however, that an evolutionary perspective suggests that such rich innate and constructivist mechanisms will only be used when it is absolutely essential to do so. (The point is, however, noticed by Tennant (1984) in his comments on Chomskian linguistics.)For evolutionary pressure should favour the lowest order of internal complexity capable of serving the needs of an organism in a given niche. Constructivist additions will be a source of possible errors in cognition. Going beyond the data always implies a certain risk, for the environment may change in such a way that the strategy no longer yields useful knowledge of the world.

My point, in short, is that JJF a being can possibly extract the relevant data from its environment instead of inferring it on a meagre data-base,—then -evolution wilJL-favour- the -development -43 f^-the-requi-red-informational sensitivity over the development of a constructivist stand-in. To this extent, the ecological movement in psychology is partially vindicated. It seems quite likely that a large part of our cognitive achievement is due to extensive and unmarked sensitivities to ambient information. To this extent, then, I would argue (see also Rutkowska (1985)) that work in A.I is often too cognitively encapsulated. The methodological solipsist strategy of formulating models of mind without considering how mind is embedded in an information rich world will, if I am right, prove another source of cognitive wheels.

Rutkowska draws on studies in developmental psychology to support a similar conclusion. Thus she notes that children use the physical structure of their environment to aid their problem solving activities. When seeking an ingredient for baking a cake, the child does not need to remember exactly where the ingredient is located in a store. Instead it simply goes to the right <u>shelf</u> and <u>looks</u> for what it needs. The external world thus stands in for a highly detailed memory store. (This example is cited by Rutkowska 1985 p.16.)

Features of the external world may thus often take the place of complex cognitive strategies. Some computer scientists now realise this.

Jim Nevins, a researcher into computer-controlled assembly, cites a nice example (reported in Michie and Johnston (1984)). Facing the problem of how to get a computer controlled machine to assemble tight fitting components, one solution is to compute vast series of feed-back loops telling the computer when it has failed to find a fit and getting it to try again in a slightly different way, The natural solution, however, is to mount the assembler arms in a way which allows them to give along two spatial axes. Once this is done, the parts simply slide into place 'just as if millions of tiny feedback adjustments to a rigid system were being continuously computed' (Michie and Johnston, 1984 p. 95).

Two points have now emerged. First, natural intelligence will, whenever possible, use ambient information rather than complex and fallible constructivist strategies. Second, it will often solve cognitively complex problems by exploiting physical features of its own body or environment. One of the secrets of intelligence, it seems, is to use whatever happens to be available as an aid to practical problem-solving. The last moral I draw from the biological metaphor is therefore that computational models of natural intelligence should attend to possible short-cuts provided by the embedding of the mind in a physical body in a highly structured external environment. They should not assume too readily that the explanation of high-level human competences demands rich constructivist mechanisms or computational acrobatics.

Conclusions.

A proponent of the ecological movement in psychology once wrote

"Ask not what's inside your head but rather what your head's inside of".

Mace (1977). Quoted in Michaels and Carello (1981).

Such a rejection of the computational approach to psychology is surely too extreme. I have tried to suggest, however, that the computational approach, if it is to illuminate natural intelligence, must reflect on the biological history of intelligent organisms. In so doing we are led to favour the microworlds of animal intelligence as the proper strategic decomposition of our computational investigations. We are led to expect fast, fallible heuristics as guides to time-critical decision making, and we are led to expect systems which exploit ambient information and use the physical structure of the world to save our computational resources. We may therefore paraphrase Mace and advise the computational psychologist

'Ask not (just) what's inside the head. Ask also how it got there and in what physical context it is going to be used.'

Andy Clark

REFERENCES.

DENNETT,D. (1964) 'Cognitive wheels} the frame problem of AX* in C. Hookwey (Ed) WINDS, MACHINES ANQ EVOLUTION (C.U.P. Cambridge, 1984)

HAYES, P. (1979) *The naive physic* manifesto* in O.fUchie (Ed) EXPERT SYSTEWS IN THE BICRO-ELECTRONIC ASE (Edinburgh Univ. Press, Edinburgh, 1979)

LORENZ, K. (1941) *Kant*s Lehre von Apriorischen in Lichte flegetwe'rtiger BiolOQle[1] BLATTER rUROEUTSCHE PHILOSOPHIE 15, 1941.

HARR, D. (1977) •Artificial Xntalligancat • personal view* in 3. Haygeland (Ed) HIND DESIGN (BIT Press 1981)

MICHAELS, C. and CARELLO, C. (1981) DIRECT PERCEPTION (Prentice-Hell.New 3ereey 1981)

MICHIE, D. and JOHNSTON, R. (1984) THE CREATIVE COWPOTER (Penguln,England, 1984)

RuTK0WSKA,3. (1985) •Developmental Psychology Contribution to Cognitive Science' (Paper presantsd to SEAKE Conference on "Artificial Intelligence for Society" Brighton Polytechnic, Duly, 1965.)

SCHfINK.R. and ABELSON.R. (1977) SCRIPTS.PUNS.GOALS AND UNDERSTANDING (Laurence Erlbaun Associates, New 3eraey, 1977)

SLOfIAN,A. (19e5) *Real time Multiple-active systems* (draft manuscript).

TENNANT, N. and SCHILCHER.F. (1984) PHILOSOPHY.EVOLUTION AND HUWAN NATURE (Routledge *nd Kegan Paul, England, 1984).

WINOGRAO, T. (1972) 'Understanding natural language* COGNITIVE PSYCHOLOGY I 1972.