NEURAL HARDWARE AND THE PRESUMED AUTONOMY OF PSYCHOLOGY

William Bechtel
Bernard Ecanow
University of Illinois Medical Center

Two types of arguments are commonly given in support of the claim that cognitive psychology can predict and explain cognitive processes without troubling itself with the details of neurophysiology. The justified conclusion of these arguments is often thought to be that artificial intelligence research, which tries to model human thought on electronic hardware, "can be regarded as psychology in a particularly pure and abstract form [since] the same fundamental parameters are under direct experimental control (in the programming), without any messy physiology or ethics to get in the way" (Haugeland, 1981, p. 31). This paper will challenge the validity of both arguments for this claim and propose how features of neurological hardware may have consequences for the performance of human cognition.

The first argument for the autonomy of psychology originated with Putnam (1975) and has been developed most extensively by Fodor (1974 and 1979). Putnam noted that in the case of computers the same programme can be run on very different types of hardware. Fodor extended this argument by noting that the same hardware can run alternative programmes. Thus, reduction of programme states to hardware states or of psychological states to neurophysiological ones is impossible. Psychology must thus remain a "special science" seeking its own explanatory scheme.

The second argument for the autonomy of psychology is also designed to establish the additional claim that programming computers is a particularly apt way to learn about human cognitive performance. This argument starts with the assumption that all the information humans can employ in their cognitive operations must cross their sensory thresholds and be encoded within them. Since it is only this encoded information that the mind-brain can employ in its information processing, Dennett describes the mind-brain as a "syntactic engine" Dennett, 1981). This argument then construes thought processes as formal processes in which one manipulates the symbols in which the information is encoded. Assuming that the mind-brain has an effective procedure for these formal processes, Church's thesis claims that there is a recursive process for computing it. Each formal process can therefore be computed by a Turing machine. Invoking the concept of a universal Turing machine (i.e., one that can imitate every specific Turing machine), the argument concludes that thought processes can be modelled on a universal Turing machine. Psychology can direct itself to producing computer or Turing machine models that replicate human thought and not concern itself with neurophysiology.

As enticing as these arguments make the prospects of an autonomous psychology seem, they are seriously flawed. As Richardson (1979) has argued, even if the mapping between neurophysiological states and psychological states is many-many, that does not eliminate the possibility of an informative reduction of psychology to neurophysiology. All that is required are neurological conditions that are sufficient for determining the psychological states. Moreover, if the argument Putnam and Fodor use against the explanatory relevance of neurophysiology works, it also undercuts the simple appeal to programming models to explain cognitive functions. Just as the same programme can be run on different hardware, different programmes can account for the same behavior. Therefore, even if a programme perfectly minics human behavior, one has no assurance that it actually describes how humans manipulate symbols (cf. Bechtel, forthcoming).

The second argument for the autonomy of psychology is just as flawed. This argument moved from claiming that symbol manipulation can be modeled on a universal Turing machine to using actual computer programmes to model that process. Haugeland notes what is assumed in making that move: "with one qualification, universal machines can be built, that is what digital computers are. This one qualification is that a true universal machine must have unlimited storage, whereas any actual machine will have only a certain fixed amount" (Haugeland, 1981, p. 13). That qualification, however, has very far reaching consequences. Neither our brains nor digital computers come close to having the unlimited resources required by a universal Turing machines. With limited resources, however, neither brains nor computers can employ the kinds of algorithms that Church's thesis assures us exist for all decidable processes. So the use of Church's thesis and the concept of a universal Turing machine to justify using computer similation as a way of studying human psychology is unjustified.

Neither of these responses to the arguments for the autonomy of psychology from neurophysiology shows that psychology might not profitably be pursued in this autonomous manner or that computer modelling might not be the most powerful means of doing that. But they do undermine the assumption that artificial intelligence models provide an adequate basis for understanding human cognition. While not denying that such models can show us interesting features about cognition, we shall now argue that there is reason to believe that significant differences exist between human cognition and computer models of it.

Limited resource capacity for problem solving dictates that one cannot always use procedures that guarantee correct results. For complex problems one must choose methods that yield correct results much of the time but are fallible. There are two fallible ways of using limited resources for dealing with problems whose optimal solution requires greater resources. One that has been studied much in recent years has been the use of heuristics (Cf. Simon, 1969). Heuristics are rules that are simpler than optimal algorithms, produce the same answers as the optimal algorithm much of the time, but that are subject to systematic errors because of the simplifications they use (Wimsatt, 1980). Tversky and Kahneman (1974) have developed an empirical research programme to discover the kinds of heuristics humans use in handling certain kinds of judgment tasks. A second way of solving the problem of limited resources is to manipulate the hardware of one's system to approximate the performance of a richer hardware system. As in the case of heuristics, a simplified hardware system that is developed to approximate a richer one may allow one to reach

correct answers much of the time, but will do so at the cost of making errors under certain conditions.

The hardware system of a Turing machine or a computer is linear and digital--information is processed by linearly transmitting and modifying information units which are perfectly distinct and so engender no ambiguity. One basis for the analogy between brains and computers is the assumption that the brain also utilizes a linear and digital processing mechanism--the neuron (von Neumann, 1958). Like the components of computers, neurons transmit electrical impulses linearly down their dendrites and axons with the action potential in the axon being comparable to a digital binary signal in a computer. (Dendritic processes allow for a spectrum of responses, but these functions have been viewed as weighting and gating functions, which are easily replicated in computer hardware.)

In addition to these neuronal processes, which seem comparable to those realized in a Turing machine or computer, though, there is another transmission mechanism in the brain. This mechanism involves a form of transmission quite different than the linear and digital transmission of neurons and may provide a means for the brain to approximate the performance that would require a far richer linear and digital mechanism. One can best appreciate this mechanism by considering earlier stages in evolution.

Long ago Hughlings Jackson (1884) insisted that to understand the function of the brain one had to attend to its evolution. The brain is organized in an evolutionary hierarchy in which the lowest and first evolved parts of the brain regulate all bodily activities. The later evolved higher centers function by modifying and regulating the earlier evolved lower centers. Before there were nerve cells, however, there existed a mechanism for transmission between cells. According to Oparin's (1965) model, cells originated when water interacted with macromolecules and electrolytes to form a more fat like substance--protoplasm. The water around the molecules becomes structured in much the same manner as occurs in jello and the whole unit behave like an oil drop with respect to the intercellular plasma. Ecanow (1982) has proposed that the same process is responsible for forming multicellular aggregates. In these aggregates different thermodynamic states are found in the cytoplasm of the various cells (including a different state in the cellular interface or membrane) and in the interstitual fluids.

Already within these early cellular aggregates a mode of signal transmission existed. The different thermodynamic states of the cytoplasm, membrane, and interstitual fluids are in dynamic equilibrium with one another, with a constant exchange of molecular substances occurring between the different structural units. This exchange allows for a kind of transmission between cells: a change in the thermodynamic conditions in one cell will propagate rapidly to surrounding cells. This kind of transmission still occurs even after nerve cells have evolved with their more digitalized and linearly directed transmission capacities. This is particularly true in places where nerve cells are tightly bound together. This tightly organized pattern causes the water in the plasma surrounding the cells to become highly structured itself, affecting, in particular, the solubility of ions in the plasma. Both the cells and the surrounding plasma become highly susceptible to any thermodynamic changes that are induced. One of the prime causes of thermodynamic changes is electrical activity propagated along neurons. Electrical energy alters the physical-chemical structure around the nerve cell. Once the change has occurred, the physical-chemical organization elsewhere will modify until equilibrium is once again achieved. Not only are these physical-chemical changes initiated by neural activity, they also reciprocally affect that activity. Neural activity depends on ion transfer, and this ion transfer is governed by the degree of structuring found at the cell-plasma interface. One cell's firing changes this structuring around other cells and hence their potential to fire.

There is, at this point, reason to believe this physical-chemical transmission mechanism is efficacious in humans. Since most anesthetic agents are biochemically inert, it is generally recognized that a physical-chemical mechanism must be involved. Following a suggestion from Bernard (1875), Ecanow et. al. (1979 and Ecanow, 1981) propose that anesthesia involves the formation of a highly structured matrix at the cell surface which becomes non-polar and fat-like. Ion exchange is a polar process and so is blocked in such a matrix. This model predicts that substances which decrease the structuring of water, generally polar molecules, chaotropic ions like urea and vitamin C, or increased temperatures, will produce an increase in mental activity. These effects have been observed in vivo. The insight of this model has been extended to account for the fluctuation between increased and reduced mental activity found in manic-depressive patients (Ecanow and Klawans, 1974).

This physical-chemical mode of transmission proposed by Ecanow (1982) differs from neural transmission in propagating three dimensionally from the initial site and in invoking a degree of response that can vary over a continuous spectrum. It is also very rapid. We cannot, at this point, make definitive claims as to its direct role in cognition, but we conclude with a speculative suggestion. Kandel (1978) has found that long term and short term habituation and sensitization in Aplysia (processes he takes to be forms of memory and learning) result from changing the amounts of calcuim ions (needed for transmitter release) available at the pre-synaptic cleft. Kandel does not account for the change in calcium availability that habituation and sensitization produce, but one possible mechanism would be through alternation of the physical-chemical structures near the pre-synaptic cleft. Such structuring can occur in degrees and so account for the gradual "learning" of these responses. Moreover, such structures would be appropriately subject to change when new experiences produce neural activity in the area around the pre-synaptic cleft.

The physical-chemical transmission mechanism provides the mind-brain with capacities for information processing quite different from the linear and digital capacities of neurons. Given the hardware limitations of the brain, it may well be that this three dimensional analogue mechanism of physical-chemical transmission provides the mind-brain a powerful tool for overcoming resource constraints. The power of this mechanism, however, cannot be studied by modelling with digital computers that lack such transmission capacities.

REFERENCES

Bechtel, William (forthcoming), "Two Common Error in Explaining Biological and Psychological Phenomena," Philosophy of Science.

Bernard, Claude (1875), Lecons sur les Anesthetiques. Paris: Bailliere.

Dennett, Daniel C. (1981), "Three Kinds of Intentional Psychology." In R. A. Healey (ed.) Reduction, Time and Reality: Studies in the Philosophy of the Natural Sciences. Cambridge: Cambridge University Press.

Ecanow, Bernard (1981), "A Comprehensive Theory of Anesthesia." Physiological Chemistry and Physics Journal 13: 23-27.

Ecanow, Bernard (1982), "Interstitial Conduction and the Emergent Mind." Journal of Pharmaceutical Sciences 71: viii.

Ecanow, Bernard and Klawans, H. L. (1974), "Physical-Chemical Properties of Cellular Constituents and their Contribution to Neuronal Function," in H. L. Klawans, (ed.) Models of Human Neurological Diseases. Amsterdam, Excerpta Medica.

Ecanow, Bernard, Gold, B. H., and Ecanow, C. S. (1979), "Unified Theory of Anesthesia," Journal of Pharmaceutical Sciences 68: iv-v.

Fodor, Jerold (1974), "Special Sciences," Synthese 28: 97-115.

Fodor, Jerold (1979), The Language of Thought, Cambridge: Harvard University Press.

Haugeland, John (1981), "Semantic Engines: An Introduction to Mind Design." In J. Haugeland, (ed.) Mind Design. Montgomery, VT: Bradford Books.

Jackson, J. Hughlings (1884), The Croonian Lectures on the Evolution and Dissolution of the Nervous System, London.

Kandel, Eric (1978), "Small Systems of Neurons," Scientific American 238: 66-76.

Oparin, A. I. (1965), "The Pathways of the Primary Development of Metabolism and Artificial Modeling of this Development in Coacervate Drugs." In S. W. Fox (ed.) Origins of Prebiological Systems and of their Molecular Matrices. New York: Academic.

Putnam, Hilary (1975), Mind, Language, and Reality: Philosophical Papers, Volume 2. Cambridge: Cambridge University Press.

Richardson, Robert C. (1979), "Functionalism and Reductionism," Philosophy of Science 46: 533-558.

Simon, Herbert A. (1969), The Science of the Artificial, Cambridge: M.I.T. Press.

Tversky, Amos and Kahneman, Daniel (1974), "Judgment Under Uncertainty: Heuristics and Biases." Science 185: 1124-1131.

von Neumann, John (1958), The Computer and the Brain, New Haven: Yale University Press.

Wimsatt, William C. (1980), "Reductionistic Research Strategies and Their Biases in the Units of Selection Controversy," in T. Nickles (ed.), Scientific Discovery: Case Studies, Dordrecht: Reidel.