The Cognitive Animal

Empirical and Theoretical Perspectives on Animal Cognition

OA Funding Provided By:

The open access edition of this book was made possible by generous funding from Arcadia—a charitable fund of Lisbet Rausing and Peter Baldwin.

The title-level DOI for this work is:

doi:10.7551/mitpress/1885.001.0001

21 From Cognition in Animals to Cognition in Superorganisms

Charles E. Taylor

Research Questions

My research in cognition has not been directed toward animal cognition per se. Rather, I have been intrigued by the apparent inability of modern science, as I understand it, to explain subjective experience. For this discussion I will refer to it as "the problem of mind." While they are not the same, there is a broad overlap between this problem and those of animal cognition.

My research on the problem of mind has been motivated by beliefs, in no way original with me, that:

- Far from consciousness being explained by anyone, I do not think we even know what an adequate explanation would look like, an opinion articulated especially well by Nagel (1986).
- Probably some animals experience subjective feelings in more or less the same way that we do, although there is a gradation.
- At present, however, we do not know how to objectively identify whether others, animal or human, experience subjective feelings or what those feelings are like, more or less for reasons discussed by Erwin Schrödinger (1967) as "objectivation."
- The study of artificial intelligence (AI) and robotics has contributed a great deal to our understanding of the problem, but certainly has not solved or explained it.

However, it does seem to me that:

- While we are not now able to construct an artificial system that can obviously experience subjective feeling, we may still be able to evolve systems that do, using techniques that have come to be known as evolutionary computation (Mitchell and Taylor 1999).

In recent years, largely as a result of work by Rodney Brooks and his students, much of which is collected in Brooks (1999), I have come to believe further that:

- For purposes of understanding cognition, especially animal cognition, and probably subjective experience, organisms are best viewed as collections of sensors, effectors, and processors of limited abilities surrounding and located throughout the organism. These collections communicate primarily with other sensors, effectors, or processes that are mostly nearby, that mostly have a limited bandwidth, and that function as an ensemble.

This means that cognition and subjective experience are to be understood as the collective behavior of these limited agents acting as one. If so, then one of the most important challenges facing us, but one we can probably address in concrete terms, will be to understand how such an ensemble can function as some sort of superorganism. In the remainder of this essay I attempt to explain how and why I arrived at this opinion, and at the end briefly address some of the promises and limitations it offers.

Past Research

Initially I approached the problem of mind by looking at animals that could reasonably be assumed to experience rudimentary subjectivity. We used mutant strains of *Drosophila melanogaster*, exploring how learning or memory-retention mutations affected habitat selection (Taylor 1987). These insects were chosen because much is known about their biology and because they are easy to breed and evolve. I soon abandoned this approach, however, because it seemed difficult for me to relate to their sensory experiences; they were setting the boundary conditions, not I. It was too difficult to broach the subjective-objective problem mentioned earlier.

Artificial systems, however, can be made more or less complex, as desired. In the early 1970s there was enthusiasm that the "traditional" approach to artificial intelligence would provide insights to the problems of cognition. By "traditional" I mean the approach exemplified by Newell and Simon (1972) in their General Problem Solver (GPS) program, for example. Very simply, a cognitive system is viewed as having a snapshot of a state space encoded by symbols. There are rules by which these symbols may be manipulated, and cognition consists of searching through a sequence of manipulations of these symbols so that they will achieve the desired configuration of state space in a suitably economical fashion. According to this view, problem solving by animals consists of obtaining information about their environment (describable by predicate calculus) and then manipulating this representation to obtain a desired state of affairs, again describable by predicate calculus. The animal would then follow that sequence. Cognition is the deliberation about which chain of actions is most appropriate.

There is a school of thought in evolutionary biology that views the environment as posing problems that animals need to solve and that sees evolution as guiding them in doing so. Pursuing this approach, I developed a computer system in which objects inherited certain propositions at the time they were created. Other propositions could be obtained from the environment. The objects would then act on these propositions using the rules of symbolic logic to see if they could solve problems posed by the environment. Those who did so could reproduce, those who were unable to arrive at those solutions perished without leaving offspring. The system was written in Lisp and worked only modestly well, so I did not publish it. Since then, others, especially John Koza (see Fogel 1998; Mitchell and Taylor 1999), have developed powerful Lisp-based systems for evolutionary computation termed *genetic programming*. Koza et al. (1992) were able to obtain some insights about (possible) reasoning by anolis lizards when the lizards choose to pursue or ignore insects while foraging. Audrey Cramer, who started her research in my laboratory as a graduate student, subsequently explored how similar models of cognition might explain how vervet monkeys choose a route in their search for food (Cramer and Gallistel 1997).

In the early 1980s I began working with David Jefferson on a system he had developed that embodies what I feel characterizes the nature of organisms and life. We have articulated this in several papers and the work is summarized in Taylor and Jefferson (1995). Most important is the view that life is a property of *processes*, rather than of *objects*. Living beings are self-organizing processes (or even an ensemble of processes) that interact with the world around them, taking in some things and expelling others, possibly reproducing other such processes, and eventually "dying," in that they are no longer capable of self-perpetuation. Sentiments about process versus object that influenced my thinking include those of Whitehead (1926) and Birch and Cobb (1981), although I have been unable to reconcile certain significant differences between us. In any event, it seems likely to me that mind as we know it can be described only by "living" processes or objects embodying such processes. If so, then my prior work with problem solving had fundamental shortcomings.

Rob Collins, a student of Jefferson's, used connection machines—computers with tens of thousands of processors, each with its own memory and processes—to evolve populations of processes that seemed lifelike in that they could reproduce, move, die, and evolve as they navigated a virtual world (Jefferson et al. 1992). Such systems could also learn, not simply evolve—a combination that has great adaptive potential (Belew and Mitchell 1996).

Jefferson and I became acquainted with the theories of Rodney Brooks, a Massachusetts Institute of Technology (MIT) exponent of "nouvelle AI." In a series of papers with titles like "Intelligence without reason" and "Elephants don't play chess" (collected in Brooks 1999), he argued that intelligent systems need to be (1) embodied, not just simulated; (2) situated in the real world, not simply in a demonstration setting; and that (3) a subsumption architecture is the best way to control such systems. Brooks's views were heavily influenced by certain ethologists, especially von Uexküll (1921). In this view, "traditional AI" is incapable of capturing animal intelligence. Among other reasons, the representation of natural environments by traditional AI models in a way that is sufficiently rich for intelligent behavior requires an unfeasible number of logical manipulations. Brooks suggested that the best way for an animal to represent its environment is to use the environment to represent itself.

To test such theories, one cannot use manipulation of disembodied symbols, but generally needs to use robots in rich environments. These robots have typically consisted of large numbers of distributed sensors, effectors, and processors, each with modest abilities. None of these processors could be said to characterize the environment in any meaningful way by itself, yet collectively these systems are capable of complex coordinated actions such as climbing up stairs or collecting empty soda cans from desk tops. Furthermore, watching robots play tag or attempt to fetch an object comes much closer to meeting the criteria that are now used to infer subjective feeling, described in the introduction, than watching computers that simply manipulate symbols. Hence we thought it worth exploring whether evolving robots might contribute to studying the problem of mind.

At that time few robots were commercially available and supported. So two students in my laboratory, Kourosh Nafisi and Orazio Miglino, constructed a small, fairly robust robot out of Lego blocks using a small controller board we obtained from the MIT artificial intelligence laboratory (Miglino et al. 1994). It resembled an early version of the Lego Technics kits that are now sold in toy stores. It was controlled by neural nets, and populations of such nets could be made to evolve by a variant of genetic algorithms (see Mitchell and Taylor 1999). Miglino's robot evolved fairly complex neural networks and the associated behavior needed to navigate a simple environment. These robots were a hybrid combining simulated neural nets, similar to those evolved by Collins (Jefferson et al. 1992), and real measurements of occasional physically realized rules. We were unable to physically test all of the robots that needed to be tested each generation because the robots were only plastic and were not sufficiently robust to withstand hours of continuous use; nor was it possible to provide power to them for fully automatic testing. To my knowledge, these were the first hybrid simulated–actual robots that were created by evolutionary computation. Since that time, very robust robots have been developed and ways have been found to provide constant power to them (Nolfi and Floreano 2001). One of these, the Khepera robot, has become the workhorse of evolutionary robotics, although even these versions are still very simple and require an environment so artificial that they can be regarded only as embodied, not situated in the real world (R. A. Brooks, personal communication).

As we attempted to reconcile these studies with those of Brooks, it became evident that a critical feature of intelligent behavior by real organisms is the ability for large numbers of sensors, processors, and actuators to function as a unified ensemble. A robot is surrounded and studded with many such elements communicating together. None of the elements has a complete model of the environment—certainly not one that could be naturally represented by propositional calculus as the earlier AI models required—but each element has some ability to digest and process the information it obtains locally and from sensors and processors. In a living organism, however, cognition seems to be almost a collective consciousness derived from the interaction of numerous parts functioning as a "superorganism."

Current Research

If we are to evolve such a "collective individual" as a robot, we will need to pay much more attention to just what is abstracted from the environment and communicated to other parts of the organism. A collection of sensors and processors indiscriminately pouring their electrons into a bus is not likely to achieve our goal. Instead, each sensor and processor needs to compress the experience that learning or evolution has judged to be important and transmit only that information. It is especially challenging to make the experiences of different sensors and processors mutually intelligible.

With a computational linguist, Edward Stabler, and a student in my laboratory, Tracy Teal, I have begun to explore how such extraction and compression might be achieved. Our first work has been exclusively with symbolic systems (Teal et al. 1999; Teal and Taylor 2000). We are now exploring how to relate such symbols to actual experiences (e.g., in Wee et al. 2001). [Several years ago I cooperated with Takaya Arita on a study in which populations of neural nets learned to acquire a common lexicon, even when they were not all able to observe the same object (Arita and Taylor 1996).]

If this approach is more or less correct, then the challenge ahead is to learn how distributed systems operating in the real world are able to combine and adapt to construct a unified whole —possibly to the extent of making a "whole iguana" (Dennett 1978). This will necessarily involve both mechanical and electrical engineering and computer science, with plenty of theory and real-world constructions. Probably linguistics and evolutionary theory will provide needed insights.

Phylogenetic Appropriateness

The overriding question or test of this approach is one of phylogenetic appropriateness. Is in silico life going to be sufficiently similar to in vivo life, in the way that much in vitro life has proven to be? Will these simpler and manipulable constructions of intelligent behavior really lead to an understanding of mind and cognition? Or will they be merely a diversion, albeit one that is enthralling and no doubt commercially useful, but still just a diversion, from our understanding of the deeper issues?

Although they are poorly understood, subjective and physical events are undeniably part of the same phenomena. "The world is given to me only once, not one existing and one perceived" (Schrödinger 1967, p. 137). Putting aside the problem of how subjective experience is to be recognized at all, some physical systems must be capable of subjective experience and some (we believe) must not be, for reasons that are not known. We are currently unable to judge just which system might become capable and which will not. If a system is capable of becoming subjective by incremental changes, and if our criteria for selection are appropriate, then it should be possible to evolve systems that have subjective experiences. But what if the system is simply not capable of the physical interactions necessary to support subjective experience? A liter of oxygen, though unquestionably a physical entity, is capable of becoming solid only under the most unusual of circumstances, if at all. It might be that the silicon chips of an artificial construct are similarly not capable of the physical processes that must underlie subjective experiences. Maybe the necessary interactions occur only between carbon, hydrogen, and oxygen. If that is the case, then computers and robots as now constructed will be phylogenetically inappropriate vehicles for studying the problem of mind.

Or maybe they will have a different sort of subjective experience. Extrapolating from the imaginative and forceful arguments of von Uexküll (1934/1957), we can presume that the "experience" of such creations will be quite different from our own, and probably quite varied from creation to creation. There is every reason to expect that these creations will provide us with an enormously useful test-bed for studying the relation between cognition and subjective experience; perhaps they may even provide evidence for other forms of conscious experience, called for by Nagel (1986).

In the past few decades we have learned a lot from efforts to create computers and robots that behave in ways that might seem to require subjectivity. One need only peruse the robots on display at http://www.androidworld.com/ to be impressed by their variety and accomplishments. Most emphasis has been on human subjectivity, but there has been progress from attempts to explore intelligent behavior and subjective experience at other levels, including that of lizards. There is no reason to believe this progress will stop, and there is plenty of reason for optimism. Whether this category of model is truly phylogenetically appropriate or not, only time will tell.

References

Arita, T. and Taylor, C. (1996). A simple model for the evolution of communication. In *Evolutionary Programming: Proceedings of the Fifth Annual Conference of Evolutionary Programming*, L. Fogel, P. Angeline, and T. Bäck, eds., pp. 405–410. Cambridge, Mass.: MIT Press.

Belew, R. K. and Mitchell, M. (1996). *Adaptive Individuals in Evolving Populations: Models and Algorithms*. Reading, Mass.: Addison Wesley.

Birch, C. and Cobb, J. B., Jr. (1981). *The Liberation of Life*. Cambridge: Cambridge University Press.

Brooks, R. A. (1999). *Cambrian Intelligence: The Early History of the New AI*. Cambridge, Mass.: MIT Press.

Cramer, A. E. and Gallistel, C. R. (1997). Vervet monkeys as travelling salesmen. *Nature* 387: 464.

Dennett, D. C. (1978). Why not the whole iguana? *Behavioral and Brain Sciences* 1: 103–104.

Fogel, D. B. (1998). *Evolutionary Computation; The Fossil Record*. New York: Institute of Electronics and Electrical Engineers.

Jefferson, D., Collins, R., Cooper, C., Dyer, M., Flowers, M., Korf, R., Taylor, C., and Wang, A. (1992). Evolution as a theme in artificial life: The genesys/tracker system. In *Artificial Life II*, C. G. Langton, C. Taylor, J. D. Farmer, and S. Rasmussen, eds., pp. 549–578. Reading, Mass.: Addison Wesley.

Koza, J. R., Rice, J. P., and Roughgarden, J. (1992). Evolution of food foraging strategies for the Caribbean anolis lizard using genetic programming. *Adaptive Behavior* 1: 47–74.

Miglino, O., Nafasi, K., and Taylor, C. E. (1994). Selection for wandering behavior in a small robot. *Artificial Life* 2: 101–116.

Mitchell, M. and Taylor, C. E. (1999). Evolutionary computation: An overview. *Annual Review of Ecology and Systematics* 30: 593–616.

Nagel, T. (1986). *The View from Nowhere*. New York: Oxford University Press.

Newell, A. and Simon, H. A. (1972). *Human Problem Solving*. Englewood Cliffs, N.J.: Prentice-Hall.

Nolfi, S. and Floreano, D. (2001). *Evolutionary Robotics: The Biology, Intelligence and Technology of Self-Organizing Machines*. Cambridge, Mass.: MIT Press.

Schrödinger, E. (1967). *What Is Life? Mind and Matter*. Cambridge: Cambridge University Press.

Taylor, C. E. (1987). Habitat selection within species of *Drosophila*: A review of experimental findings. *Evolutionary Ecology* 1: 389–400.

Taylor, C. E. and Jefferson, D. R. (1995). Artificial life as a tool for biological inquiry. In *Artificial Life: An Overview*, C. G. Langton, ed., pp. 1–13. Cambridge, Mass.: MIT Press.

Teal, T. and Taylor, C. E. (2000). Effects of compression on language evolution. *Artificial Life* 6: 129–143.

Teal, T., Albro, D., Stabler, E., and Taylor, C. (1999). Compression and adaptation. In *Advances in Artificial Life: Fifth European Conference on Artificial Life*, D. Flareans, J.-D. Nicoud, and F. Mordaa, eds., pp. 709–719. Heidelberg: Springer.

Uexküll, J. von (1921). *Umwelt und Innenwelt der Tiere*. Berlin: Jena.

Uexküll, J. von (1934/1957). A stroll through the worlds of animals and men. In *Instinctive Behavior: The Development of a Modern Concept*, C. H. Schiller, ed. (1957), pp. 5–80. New York: International Universities Press.

Wee, K., Collier, T., Kobele, G., Stabler, E., and Taylor, C. (2001). Natural language interface to an intrusion detection system. In *Proceedings of the International Conference on Control, Automation and Systems,* in press.

Whitehead, A. N. (1926). *Science in the Modern World.* New York, NY: Macmillan.