The Design of Learning for an Artifact

Joanna Bryson
Intelligent Systems Laboratory
Edinburgh University
Level 8, Appleton Tower
Edinburgh EH8 9JZ, UK
[44] (131) 650 4354
`joanna@ai.mit.edu`

March 8, 1996

Introduction

Why do we conduct research on learning in robotics? Presumably, in keeping with the two aims of Artificial Intelligence, either to improve the utility of our robots, or to explore models of learning in nature. Most research thus far has concentrated on developing single techniques for learning or in comparing the relative merits of these techniques. Natural intelligence, however, demonstrates the use of many different, specialized learning techniques. As McClelland, McNaughton & O'Reilly (1995) recently pointed out, research in artificial intelligence also demonstrates the utility of this approach.

This paper is a brief essay intended to help roboticists think about how exactly learning should be incorporated in their artifact. The assumption of the roboticists' goal here is that we are not building to prove a particular point or to explore a particular technique, but rather we are trying to make our robot behave intelligently. For my definition of "behave intelligently" I resort to the Turing criteria: I assume we are trying to build something that behaves like a human. In considering this goal, I propose a taxonomy of learning that I hope is useful in thinking about more general robotic tasks.

Learning: a Definition and a Taxonomy

Perhaps part of the reason we find robotic learning so difficult is we haven't been able to properly define its role for our agents. To help with my argument, I am going to work from a very general definition of learning: *learning is any process by which an agent acquires new behavior.* This is deliberately broad enough to include things like species-level learning for animals.

Consider if we were trying to build a humanoid robot, like Cog, a project under development at MIT (Brooks & Stein 1993). Let's say we began by building the body, and had only progressed to the stage where we had an aluminium skeleton, connected together by motors. At this point, does our robot embody any learning? My answer is "yes." For example, if the robot is balanced upright and I push it, it will fall, and falling is a behavior (so is balancing.) Even at this point, with no processing at all, the motions of the robot are immensely constrained. The constraints are the nature and number of joints, and the length and weight of the skeletal parts, etc. For true infants, as Brooks & Stein (1993) point out citing Lakoff & Johnson (1980), much of their comprehension of the world may be rooted in early experiences of sensing and manipulation that are directly affected by our physical bodies.

Learning such as the physical architecture of a robots body I will refer to as **provided learning.** Provided learning is any behavior the agent itself does not need to acquire. In an animal, this comes from evolution, genetic and perhaps cultural. But I choose my terminology to emphasize the role of the robot designer. Early behavior-based robots, such as Herbert and Genghis (Brooks 1990) operated entirely on provided learning; their designers built everything that determined their interactions with the environment in either software or hardware. This is still one of the primary attributes of behavior-based control; the individual behaviors tend to be heavily engineered, even though the global action sequencing is unspecified and reactive.

The rest of learning, that which *is* conducted by the agent, I divide into two classes, again emphasizing the roboticist's perspective. The first, **required learning**, or as Minsky (1986) calls it, pre-ordained learning, is prescribed behavior acquisition, often in the form of parameter setting. In robotics, archetypal examples are training periods for neural networks before full operation, or rotating after start-up to calibrate a compass to light cues. Similar learning phases occur in animals. For example, young barn owls calibrate their hearing with their vision during a certain 2-week stage of their development, after their heads' shape and thus precise acoustic properties have been determined (Gallistel, Brown, Carey, Gelman & Keil 1991). Bees learn the ephemeris of the sun for navigation; female mice imprint the smell of their impregnator on copulation. These periods of specialized plasticity in animals tend to utilize implicit learning systems and to be triggered by specific hormonal or behavioral contexts.

The final class of learning I call **open learning.** This is the general learning often considered necessary for "lifetime learning" or "true intelligence." Examples in robotics include logic or case-based reasoning systems that are now often worked in hybrid on top of behavior-based controllers. Early reinforcement learning robots attempting to derive their own control functions were almost entirely open learning. Often simulations, their only elements of provided learning were their array of possible actions, and possibly their sensing of state. Open learning covers what in natural intelligence is associative and declarative learning abilities.

Despite the advances we have made in robotic control due to Society-of-Mind type approaches, I think that as AI researchers we still tend to look for the solution to open learning in terms of overly unbiased, domain-independent systems. This is not supported by natural intelligence research. For example, classical conditioning, once considered the basis of all general learning, is actually domain-specific. A pigeon can be conditioned to peck for food, but cannot learn to peck to avoid a shock. They can, however, learn to flap their wings to avoid a shock, but not for food. Similarly, rats rapidly associate poisoned water with a smell or taste (often on one exposure) but cannot learn to associate it with visual or auditory cues, while the inverse is true for avoiding shock (Gallistel et al. 1991).

As McClelland et al. (1995) point out, trends in both natural and artificial intelligence research indicate that specialized, multi-part systems may be necessary for acquiring even strictly declarative knowledge. In an interaction similar to the role of the Baldwin effect in species evolution, a slow but reliable system learns gradually while a fast, extremely plastic, but less reliable system provides repeated exposure to the learning scenario. Expertise seems to be at its heart a massive look-up table of stored patterns; its speed of access requires that it not be a constructive process. But memorizing these complex patterns in a distributed framework requires repeated presentations — how can the situation be replicated if the skills involved haven't yet been learned? The answer is a second, separate system that is far more plastic, though less reliable and efficient. Current thinking in neurophysiology is that for mammals, the memory for these systems may be stored in the neocortex and the hippocampus respectively.

Even our apparently rational behavior, such as the ability to order objects, does not appear to be based on uniform, logic-based reasoning, but rather on specialized representational structures. Children as old as seven are incapable of ordering 7 blocks; this is a well known Piagetian task, seriation. McGonigle & Chalmers (1996) demonstrate that the underlying strategy is not "logical", that is, based on universal structuring, but that there biased grammars develop for tasks like seriation and transitivity. Further, adult monkeys, given enough time and practice, will develop structures that solve these tasks with similar response times to those in children old enough to perform them. McGonigle & Chalmers (1996) suggest that "a wide spectrum of high-level cognitive behaviours have hardwired underpinnings as necessary preconditions for their (ultimate) expression." (p. 305)

Design: a Recommendation and an Example

What does this imply for roboticists? Gallistel et al. (1991) suggest that nature always favors reliability over plasticity, that learning is used as a last resort for problems that can't be solved on an evolutionary time scale. The behavior-based AI community, with its bias towards bottom-up design, has naturally tended to follow this direction. But further, when fixed behaviors and parameter tweaking have reached their limits, learning should still be maximally restricted, modular and biased with structured frameworks.

Let's consider again the case of designing humanoid intelligence. One behavioral attribute we often assume is acquired through social reinforcement is imitation. But notice that if we *provide* imitation as a standard behavior, then we can reduce a vast array of knowledge, that already acquired by other members of our society, from open to required learning. Imitation has already been shown to be at least to some extent innate in humans; infants will imitate facial expressions of adults within their first weeks of life.

In animals, many species-typical behaviors, encoded as fixed action patterns, have both instinctive and acquired components for their triggering. A good example are the emotional responses, such as fear. We have a dedicated brain organ, the amygdala, for "switching" our emotional responses; cells fire in response to sensory input whose axons go to the other old-brain centers where the action patterns are stored. The amygdala can learn (by classical conditioning) to associate emotional responses with new stimuli. This includes both simple sensory stimuli from the older perceptual regions of the brain, such as tones or lights, and complex processed stimuli from the neocortex, such as particular rooms or individuals.

For a humanoid robot, such a module would be useful for a reward system. In order for any associative learning to take place, there must be some provided behavior for accepting reinforcement. We would want to build in social rewards, such as discriminating a smiling or frowning face, a kind or angry voice, in order to make the robot prosocial and trainable. We would also want our robot to be able to associate rewards with consistently rewarded behaviors. It could then "entertain itself" and develop skills with, say, blocks, without requiring the presence or at least the forced adulation of supervisors. The choice of the provided rewards affects the sorts of rewards our robot would learn, they would determine the "personality" of the robot. We might also want to build in certain other traits, such as a negative reaction to pressure on touch sensors over a certain weight, in order to bias it against motions that might damage itself or others, or a positive reaction to motion or bright colors in its field of view, which would base it towards situations likely to be data rich, that is, educational.

Modularity does not disallow re-use of acquired behavior in different domains, in fact this kind of re-use is crucial and apparent in nature. For example, Lorenz's imprinted male ducklings would not only follow the object they saw in their critical imprinting period as their mother, but two years later would attempt to mate with something that looked similar. For an example of how we could leverage this in our humanoid design, I'll rely on the fact that the easiest way to generalize is to fail to perceive differences.

Let's assume our humanoid has been built to simulate infancy, and one of its modules is now trying to master "conscious" control of its own (distributed) intelligence. This involves the Piaget's first task: the robot discriminating itself from its environment. One strategy would be for the module to associate acts of activation on its own part with the achievement of some goal, as recognized and rewarded by the (separate) reward module. But in the early stages, particularly if motor and perceptual skills are also being acquired, goals are most reliably met indirectly; it is easier to signal a desire to a caregiver than to fulfill a desire oneself. Coordinating activities between the different modules within the agent can be viewed as a multi-agent or cooperative problem — our robot might use the same modules to learn about controlling and coordinating its own actions as its social and perhaps physical interactions with the real world. However, there may be a trade-off for such an efficiency; our robot might experience identity problems with respect to its primary caretakers.

Conclusion

In this paper I have advocated an approach to robotic learning that is distributed and heterogeneous. This approach complements and can even be taken as a description of the behavior based strategy of control such as has been popularized by Brooks (1990). I have proposed a taxonomy of robotic behavior acquisition into provided, required and open learning, and then suggested that, following nature, learning should be designed into a system to be as limited as possible, in order for that system to perform reliably. In the history of natural intelligence research, determining the sources of behavior and the extent of its plasticity has been extremely difficult. In robots these are easier to determine (though still difficult to debug) since we know the underlying mechanisms we have built in. The ambiguity is reintroduced in the design decisions as to which behaviors should fall into which class of learning. As behavioral science and neuropsychology continue to work on the true mechanisms of natural intelligence, we can hope their discoveries will help us accomplish our goals of building intelligent robots.

Acknowledgements

Many of the ideas in this paper were developed while I was working on the Cog project, and were particularly influenced by the approach and reading suggestions from the project heads, Lynn Stein and Rod Brooks. Thanks to Will Lowe for his comments and revisions.

References

Brooks, R. A. (1990), Elephants Don't Play Chess, *in* P. Maes, ed., 'Designing Autonomous Agents: Theory and Practice from Biology to Engineering and Back', MIT Press, Cambridge, MA, pp. 3–15.

Brooks, R. A. & Stein, L. A. (1993), Building Brains for Bodies, Memo 1439, Massachusetts Institute of TechnologyArtificial Intelligence Lab, Cambridge, MA.

Gallistel, C., Brown, A. L., Carey, S., Gelman, R. & Keil, F. C. (1991), Lessons From Animal Learning for the Study of Cognitive Development, *in* S. Carey & R. Gelman, eds, 'The Epigenesis of Mind', Lawrence Erlbaum, Hillsdale, NJ.

Lakoff, G. & Johnson, M. (1980), *Metaphors We Live By*, University of Chicago Press, Chicago, Illinois.

McClelland, J. L., McNaughton, B. L. & O'Reilly, R. C. (1995), 'Why there are complementary learning systems in the hippocampus and neocortex: Insights from the successes and failures of connectionst models of learning and memory', *Psychological Review* **102**(3), 419–457.

McGonigle, B. & Chalmers, M. (1996), The Ontology of Order, *in* L. Smith, ed., 'Piaget: A Critical Assessment', Routledge.

Minsky, M. (1986), *The Society of Mind*, Simon and Schuster, New York, NY.