Learning Robots

Tom M. Mitchell
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15212 USA

Tom.Mitchell@cs.cmu.edu

August, 1990

Extended Abstract[1]

How might one organize robots so that they learn from their experience? This question is one of the central motivating questions underlying the study of AI. We consider this question, describe three existing robot systems that exhibit different types of learning, then consider the open issues which remain to be addressed.

Consider a robot to be any system which interfaces to an external world via some sensors (e.g., video camera) and effectors (e.g., a robot arm). The *robot problem* is for the robot to choose, based on it sensor inputs, effector actions that affect its world to achieve its goals (i.e., desired world states). The *learning robot problem* is to construct the robot in such a way that it automatically becomes increasingly successful at this task through experience.

We consider here robots which possess the ability to predict (perhaps imperfectly) effects of their actions, and which select actions consistent with these predictions (i.e., they select actions which they predict will achieve their goals). In order for such a system to become increasingly successful with experience, it will have to exhibit several types of learning:

- It must become *increasingly correct* at predicting the effects of its actions in the world.
- It must become *increasingly reactive*, by reducing the time required for it to make rational choices; that is, the time required to choose actions consistent with these predictions and its goals.
- It must become *increasingly perceptive* at distinguishing those features of its world that impact its success.

Three systems will be described which exhibit these three types of learning.

Becoming increasingly correct. This system [Christiansen, et al, 1990] is based on a robot arm which learns to manipulate objects in a tray by tilting the tray so that the objects slide around. It begins with no theory for predicting the effects of different tray tilting actions on the positions and orientations of objects in the tray. It learns such a theory by tilting the tray and observing the consequences using its vision

---

[1] Abstract for invited presentation, To appear in *Proceedings of the 4th Australian Joint Conference on Artificial Intelligence*, Perth, November 1990.

system. This research has focused on learning models of non-deterministic actions (many of the tray tilting actions have non-deterministic outcomes since objects slide and bounce in unrepeatable ways), and different experimentation strategies for generating training data. The system learns well enough that it becomes able to successfully position and orient a rectangle within the tray with a success rate better than 95%.

Becoming increasingly reactive. This system [Mitchell, 1990] is based on a mobile robot with sonar sensors which learns to reduce its response time (the time required to react to a change in its external world) from minutes to subsecond. This robot continually senses its world and recomputes the action it believes it should perform. When possible, it computes the appropriate action by applying a set of stimulus-response rules that map directly from observed world features to the recommended action. If it has no such rules that cover the currently sensed world, it instead falls back on a traditional search-based planner that searches for a sequence of actions it predicts will achieve its goal. The system learns by applying explanation-based learning whenever it is forced to search for a plan. It thereby creates rules that recommend precisely the same action as recommended by the slower planner, in exactly those situations in which the same plan rationale would apply. However, these learned rules infer the desired action immediately from the input sense data in a single inference step--without considering explicitly the robot's goals, available actions, or their predicted consequences. Beginning with a planner, but no stimulus-response rules, this system has learned sets of stimulus-response rules which lead to rapid reactions consistent with a variety of robot goals.

Becoming increasingly perceptive. This system [Tan & Schlimmer, 1990] is based on a mobile robot with sonar sensors, which learns how to approach, sense, and identify new objects. The robot possesses a set of sensing routines that observe objects by collecting sequences of sonar readings (e.g., measure the angular width of the object by taking sonar readings as the sonar is swept from side to side). A set of training objects (e.g., a variety of cups, boxes, etc.) is placed on the floor and the classification of each is provided to the robot. The robot then collects training data by moving about the room and observing each training object using its sensing routines. From this training data it inductively infers a decision tree for sensing and classifying subsequent objects. Each test in the decision tree corresponds to applying a single sensing routine from a particular vantage point. The tip nodes of the tree correspond to the classification of the object. The learning method is able to take into account the differing costs (in seconds) of applying the different sensing routines from different locations. Thus, it attempts to construct decision trees which maximize classification accuracy while minimizing sensing cost. This system has successfully learned strategies for sensing and classifying a variety of simple, unoccluded objects, with better than 90% success rate.

While these three systems exhibit interesting capabilities in their particular domains, they constitute only the first steps toward understanding how to design learning robots. Many open research issues remain to be solved regarding each of these types of robot learning, as well as architectures for integrating these into a single system.

References:

[Christiansen, et al, 1990] A. Christiansen, M. Mason, and T.M. Mitchell, "Learning Reliable Manipulation Strategies without Initial Physical Models", *1990 IEEE International Conference on Robotics and Automation*, May, 1990.

[Mitchell, 1990] T. M. Mitchell, "Becoming Increasingly Reactive", *1990 National Conference on AI*, Cambridge, MA, August, 1990.

[Tan & Schlimmer, 1990] M. Tan and J. C. Schlimmer, "Two Case Studies in Cost-Sensitive Concept Acquisition", *1990 National Conference on AI*, Cambridge, MA, August, 1990.