Learning Distributed, Searchable, Internal Models

Dr. Richard S. Sutton
GTE Labs, Waltham, MA 02254

Although searching an internal model of the world is a standard planning technique, how such searchable world models can be learned is poorly understood. We have been taking a highly-distributed, connectionist approach to this problem: The learning system consists of a large number of individual learning elements, each of which takes responsibility for learning about a different aspect of the external world. For example, for each state descriptor there might be an element devoted to learning when a state that meets that descriptor will occur, and for each possible operator there might be an element devoted to learning when that operator is likely to be applied. Typically, many more elements are used, corresponding to various combinations of descriptors or combinations of operators and descriptors.

The representation problem implied in the selection of descriptors, operators, their combinations is being addressed by others within the connectionist framework (Rumelhart, Hinton, and Williams, 1985; Barto and Anderson, 1985; Anderson, in preparation; Ackley, Hinton, and Sejnowski, 1985). Here we assume a reasonable representation is given a priori and concentrate instead on how the cause and effect relations between elements of that representation are learned—our system learns the results of applying operators to states, what states must follow others, etc. The system also learns about its own search process—which operators are likely to be selected and which descriptors are likely to become true given the current operator selection tendencies. This knowledge is used both to evaluate possible operator selections and to direct the search process in the directions most likely to be profitable (cf. Mitchell et al., 1983; Langley, 1985; Silver, in press).

The models formed by our learning system are probabilistic, real-time, and empirical. For example, suppose it is observed that whenever the operator DROP-OBJECT is applied, a loud NOISE is heard a short time later (as the object stikes the floor). If the external world is incompletely understood, as we must assume it is when discussing learning sytems, then there will be many ways in which predictions of such consequences can be uncertain—different objects may be dropped from different heights and onto different surfaces. Perhaps a table will intervene between object and floor, perhaps the object will be a feather and make no sound, perhaps the object was never successfully picked up in the first place. The point is that it is very hard to maintain a deterministic model of the world without insisting on very high reliability and repeatability of a small set of experiences. Our approach has instead been to form explicitly probabilistic predictive models. Rather than predicting particular-sized NOISE's with particular probabilities at particular times, our models predict the expected value of the descriptor—in this case, the NOISE—within a span of time. By mixing together the size, probability, and time of occurrence of the predicted event, great simplifications in the form of the model are possible, which in turn allow much larger and more complex models to be considered.

Our current model-learning system extends that presented by Sutton and Pinette (1985) in that it permits operator-descriptor predictions as well as descriptor-descriptor predictions, and in integrating the model-search and model-learning processes. In addition, the new architecture allows conjunctions of operators and descriptors on the left-hand-side's of predictions by using the technique known as "coarse-coding" (Hinton, 1984).

We have applied these methods to computer-simulated microworlds such as maze problems. The results indicate that useful, searchable models can be learned in this distributed fashion. More-ambitious applications are currently being considered in the domain of telecommunications network management (Frawley, 1985). In these applications the system is doubly distributed in that overall network management is a function of the weakly coordinated action of separate, intelligent, model-learning and planning systems at each node, each of which is itself a distributed, connectionist system.

A larger objective of this line of research is to explore the extent to which connectionist systems and heuristic search are compatible. Connectionist structures seem to be well suited to learning methods, particularly in uncertain and noisy domains. As such, they are an attractive direction in which to look for ways of learning searchable models. In any event, connectionist heuristic-search systems should allow a better understanding of the relative strengths and weaknesses of connectionist and more-traditional symbolic approaches.

REFERENCES

Ackley, D.H., Hinton, G.E., and Sejnowski, T.J. (1985) "A learning algorithm for Boltzmann machines," *Cognitive Science, 9*, 147–169.

Anderson, C.W. Learning new terms and representational bias in connectionist systems, Ph.D. Dissertation, in preparation.

Barto, A.G. and Anderson, C.W. (1985) "Structural learning in connectionist systems," *Proceeding of the Seventh Annual Conference of the Cognitive Science Society*, 44–53, August, 1985.

Frawley, W.J. (1985) Artificial intelligence and knowledge based systems research to support future telecommunications. Technical Report 85-178.1, GTE Labs, Waltham, MA 02254.

Hinton, G.E. (1984) Distributed representations. Technical Report CMU-CS-84-157, Carnegie-Mellon Univ, Pittsburgh.

Langley, P. (1985) "Learning to search: From weak methods to domain-specific heuristics," *Cognitive Science, 9*, 217–260.

Mitchell, T.M., Utgoff, P.E., and Banerji, R. (1983) "Learning by experimentation: Acquiring and refining problem solving heuristics," in R.S. Michalski, J.G. Carbonell, and T.M. Mitchell (Eds.) *Machine Learning: An Artificial Intelligence Approach*, Tioga.

Rumelhart, D.E., Hinton, G.E., and Williams, R.J. (1985) Learning internal representations by error propagation. Technical Report 8506, Institute for Cognitive Science, UCSD, 1985. Also in: D.E. Rumelhart and J.A. McClelland (Eds.) *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, (in press).

Silver, B. *Meta-Level Inference: Representing and Learning Control Information*, North-Holland, in press.

Sutton, R.S. and Pinette, B. "The learning of world models by connectionist networks," *Proceeding of the Seventh Annual Conference of the Cognitive Science Society*, 54–64, August, 1985.