Dissertation Abstract

PAGODA: A Model for Autonomous Learning in Probabilistic Domains

Marie desJardins

My Ph.D. dissertation (desJardins 1992a)[1] describes PAGODA (probabilistic autonomous goal-directed agent), a model for an intelligent agent that learns autonomously in domains containing uncertainty. The ultimate goal of this line of research is to develop intelligent problem-solving and planning systems that operate in complex domains, largely function autonomously, use whatever knowledge is available to them, and learn from their experience. PAGODA was motivated by two specific requirements: The agent should be capable of operating with minimal intervention from humans, and it should be able to cope with uncertainty (which can be the result of inaccurate sensors, a nondeterministic environment, complexity, or sensory limitations). I argue that the principles of probability theory and decision theory can be used to build rational agents that satisfy these requirements.

PAGODA incorporates innovative techniques for using the agent's existing knowledge to guide and constrain the learning process as well as a powerful new mechanism for representing, reasoning with, and learning probabilistic knowledge. Additionally, PAGODA provides a conceptual framework for addressing important open problems such as incremental, resource-bounded learning and knowledge-based learning and planning.

PAGODA was implemented in the RALPH (rational agent with limited processing hardware) world, a simulated two-dimensional world used at the University of California at Berkeley as a test bed for designing intelligent agents. The system consists of an overall agent architecture and five components within the architecture. The five components are (1) goal-directed learning (GDL), a decision-theoretic method for selecting learning goals; (2) probabilistic bias evaluation (PBE), a technique for using probabilistic background knowledge to select learning biases for the learning goals; (3) uniquely predictive theories (UPTs) and probability computation using independence (PCI), a probabilistic representation and Bayesian inference method for the agent's theories; (4) a probabilistic learning component, consisting of a heuristic search algorithm and a Bayesian method for evaluating proposed theories; and (5) a decision-theoretic probabilistic planner, which searches through the probability space defined by the agent's current theory to select the best action. PAGODA is given as input an initial planning goal (its overall behavioral goal, for example, "maximize utility") and probabilistic background knowledge about the domain. The agent selects learning goals (features in the world about which predictive theories will be formed) that will maximize the agent's ability to achieve its planning goal. The learner uses the probabilistic background knowledge to select learning biases for each learning goal and performs a heuristic search through the space of UPTs, using Bayesian techniques to evaluate the theories. The planner uses the best current theory to choose actions that satisfy its planning goal and to generate new learning goals for the inductive learning subsystem to focus on.

Goal-Directed Learning

GDL is a decision-theoretic method that PAGODA uses to decide which learning goals will be most effective at increasing its ability to achieve its planning goal (desJardins 1992b). PAGODA's initial learning goal is just its planning goal; that is, it initially learns a theory that predicts the value of its planning goal at each time step. GDL then uses this theory to determine the degree to which features in the sensory input affect the outcomes of its actions. The features that have the greatest effect are selected as learning goals. For example, in the RALPH world, the :MUHCH action (which RALPH uses to eat food) has no effect unless RALPH is at the same location as a food object. Learning to predict what actions move RALPH to a location containing food enables the agent to form better long-term plans; therefore, "being at food" is a useful learning goal.

Probabilistic Bias Evaluation

An autonomous agent must be able to select biases (Mitchell 1980) for new learning tasks as they arise. PBE uses probabilistic background knowledge and a model of the system's expected learning performance to compute the expected value of learning biases for each learning goal. PAGODA uses the bias with the highest expected value for learning.

The probabilistic background knowledge consists of a set of uniformities, which are a probabilistic version of determinations (Davies and Russell 1987). A uniformity specifies the degree of relevance of a set of features F for predicting an outcome O. The expected asymptotic accuracy that can be achieved using a proposed bias (the feature set F) to predict a learning goal (the outcome O) is derived from the uniformities. Accuracy as a function of time (number of observations made) is found by applying a prespecified learning curve to this asymptotic accuracy. Finally, the overall predictive accuracy of the bias is found by discounting the accuracy, that is, by

multiplying the accuracy by a discounting function and integrating over time. The discounting function models the effect of the passage of time on the value of predictions. The resulting expected discounted future accuracy is used as the expected value of the bias.

Representing Probabilistic Theories

PAGODA's learned knowledge is represented as probabilistic theories about its learning goals. Each theory consists of a set of conditional probabilities predicting the value of a learning goal, or feature of the world. PAGODA's theories are called UPTs: Their form is constrained in such a way that a straightforward set of independence assumptions can be applied to the theory to make a unique probabilistic prediction about the learning goal whose value is predicted by the theory given any input (that is, any value for the agent's sensory input and any action taken by the agent). The predictions are made by the PCI inference method. PCI finds the set of most-specific conditional probabilities in a theory that apply to a given input. These probabilities are combined by iteratively finding a separable rule in the set,[2] computing its contribution to the overall probabilistic prediction using independence assumptions, and recursively processing the remaining probabilities.

Probabilistic Learning

The learning component utilizes splitting and merging operations (similar to those used in COBWEB [Fisher 1987]) to search the space of theories. The learning process is incremental in that it processes each training example separately, yielding a set of best theories after each one; however, because all past training instances are stored and potentially reprocessed, the learning process is not bounded in time or space. Proposed theories are evaluated using a Bayesian evaluation function, where the prior probability of a theory depends on its simplicity, and its likelihood depends on its predictive

accuracy on the set of training examples seen so far. The simplicity of a theory is based on the minimum description length principle (Rissanen 1978) but can vary depending on the encoding scheme used. In the dissertation, four different encoding schemes are discussed, corresponding to four levels of abstraction for classifying theories.

Probabilistic Planning

PAGODA's planning method is based on decision theory (von Neumann and Morgenstern 1947). The planner performs a random action a fixed percentage of the time (the default is 25 percent) to ensure that the environment is continuously explored and that the agent does not get stuck on a local maximum. The rest of the time, the planner performs a heuristic search through the space of possible outcomes (as defined by the agent's current best theory) for each action; this search extends to a fixed depth (default 3). The maximum expected utilities in each final state are propagated backward, yielding an expected utility for each initial action. At each level, the utility of the action with the highest expected (average) utility is used as the value to propagate back.

Results

PAGODA was tested in the RALPH world. The results of these tests show that PAGODA learns relatively accurate (that is, significantly better than chance) probabilistic models of the world and that using these models allows the system to improve its overall performance. However, the tests also highlight areas for future research, particularly improving the heuristic search for theories and the probabilistic planning mechanism.

Notes

1. This article is available as Technical Report UCB/CSD 92/678 from the Department of Computer Science, University of California at Berkeley, Berkeley, CA 94720.

2. A rule is separable within a set if its conditioning context can be split into a set of features that are shared with no other rule in the set and another set of features that are shared with some single other rule in the set.

References

Davies, T., and Russell, S. 1987. A Logical Approach to Reasoning by Analogy, Technical Report 385, AI Center, SRI International, Menlo Park, California.

desJardins, M. 1992a. Goal-Directed Learning: A Decision-Theoretic Model for Deciding What to Learn Next. Paper presented at the 1992 Machine Discovery Workshop, Aberdeen, Scotland, 4 July.

desJardins, M. E. 1992b. PAGODA: A Model for Autonomous Learning in Probabilistic Domains. Ph.D. diss., Dept. of Computer Science, Univ. of California at Berkeley.

Fisher, D. H. 1987. Knowledge Acquisition via Incremental Conceptual Clustering. *Machine Learning* 2:139–172.

Mitchell, T. 1980. The Need for Biases in Learning Generalizations, Technical Report, CBM-TR-117, Dept. of Computer Science, Rutgers Univ.

Rissanen, J. 1978. Modeling by Shortest Data Description. *Automatica* 14:465–471.

von Neumann, J., and Morgenstern, O. 1947. *Theory of Games and Economic Behavior.* Princeton, N.J.: Princeton University Press.

Marie desJardins is a member of the Applied Artificial Intelligence Program at SRI International, Menlo Park, California. She received her Ph.D. in computer science from the University of California at Berkeley in 1992. Her current interests include machine learning, planning and scheduling, temporal reasoning, and intelligent tutoring systems.