Learning and Planning in Huge Uncertain Environments

Terran Lane & Leslie Pack Kaelbling

Artificial Intelligence Laboratory
Massachusetts Institue Of Technology
Cambridge, Massachusetts 02139

http://www.ai.mit.edu

The Problem: We want to build a robust learning and planning system for robots, in which:

- A huge model of the dynamics of the world is acquired by learning small pieces at a time, represented as Bayesian networks, and then connecting them together.
- Planning is performed in an adaptive hierarchy. At the most abstract levels, it will plan for long term objectives in small models that capture only the relevant attributes of the huge model. For instance, in trying to plan to drive home from work, I might construct a model that involves local traffic and weather conditions and what we're going to have for dinner; the model does not need to involve the color of my car or the work I am going to do tomorrow. At lower, more concrete levels, the actions will be more "canned" routines that have been learned with model-free reinforcement methods and/or through caching of plans that were made frequently at higher levels.
- Information-gathering strategies are generated, also based on the current small, relevant models, that will enable effective achievement of goals.

Motivation: We wish to build robots that operate flexibly in very complex environments with multiple, conflicting, high-level goals, for extended periods of time. A general delivery or reconnaissance robot will live in the field for months, will traverse different kinds of terrain, will have to stop what it's doing at the moment and help with other tasks. Unfortunately, but not surprisingly, the exact solution of the general versions of the problems we are interested in is computationally intractable. We cannot hope for optimal behavior in complex environments. None of the existing techniques can even begin to address a problem of this scope.

Previous Work: Much of our [2, 3, 4] and others' research in these areas has been in the development of approximation methods that will allow the solution of larger problems. As a result of this work, we can solve problems of modest size: robust map learning and navigation in an office building, Bayesian network inference in networks of 1000 variables, behavior learning for control of 6-dimensional robots, a world-champion Backgammon player. The next major step is to solve *much* bigger problems.

Approach: We propose to construct a hierarchical version of the Plexus [1] system for "deliberative" planning in response to dynamic goals. At all times, the robot will have plans at multiple levels of abstraction. The higher the level of abstraction, the longer the horizon of the plan; the lower the level of abstraction, the more detailed the plan.

Another way to think of this is to imagine the robot with a variable-resolution view of its environment, as shown in figure 1. Parts of the world that are "close" to the robot, because they are likely to be encountered in the near future are considered at a very fine grain. Other parts of the world, that are either very unlikely to occur or that will be relevant in the future are considered much more abstractly. Thus, the robot might simultaneously

- choose its current motor actions based on a very fine view of the obstacles immediately present;
- decide whether to go to the charging station based on a more abstract view of the amount of fuel it is likely to use on the remainder of this surveillance mission; and
- decide which surveillance mission to do next based on much more abstract assessment of the enemy's likely motion.

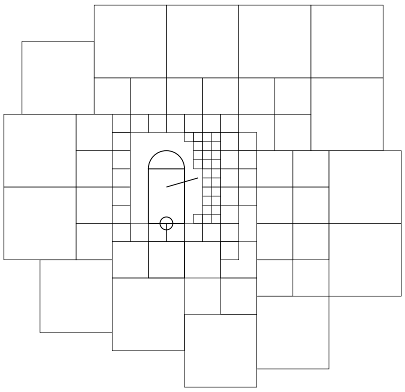

Figure 1: A variable-resolution view of the environment, in which closer, more-likely situations are considered in detail, and future, and less-likely situations are considered abstractly.

There are two motivations for a design of this sort. First, in many cases, there just isn't enough information available to plan future courses of action in detail. There is no point in thinking about what path you will take across a room that you have never encountered before; you just have to believe that you can find such a path when you get there, and have a rough estimate of how long it's likely to take. Even if complete information were available, however, complete planning at the most detailed level is almost inevitably intractable. It is practical to give up a certain amount of reinforcement in exchange for a huge computational savings.

In order to represent states at variable levels of resolution, it will be necessary to move to factored representations of states, in terms of state variables. Then, the state-transition distribution can be represented compactly by a dynamic Bayesian network.

The Plexus system could be thought of as an extremely simple instance of this architecture: states are either considered in complete detail, or are lumped into one large abstract state. Given our experience with this system, as well as with the automatic generation of and planning in hierarchical models, and with learning Bayesian network representations, we believe we can design an effective architecture along these lines.

Impact: This work may finally make it possible for robots to be deployed on complex, long-term missions with robust behavior even in very uncertain domains.

Research Support: This work is supported by DARPA/ITO under contract number DABT 63-99-1-0012.

References:

[1] Thomas Dean, Leslie Pack Kaelbling, Jak Kirman, and Ann Nicholson. Planning under time constraints in stochastic domains. *Artificial Intelligence*, 76, 1995.

[2] Milos Hauskrecht, Nicolas Meuleau, Craig Boutilier, Leslie Pack Kaelbling, and Thomas Dean. Hierarchical solution of Markov decision processes using macroactions. In *Proceedings of the Fourteenth Annual Conference on Uncertainty in Artificial Intelligence*, Madison, Wisconsin, 1998.

[3] Michael L. Littman, Anthony R. Cassandra, and Leslie Pack Kaelbling. Learning policies for partially observable environments: Scaling up. In *Proceedings of the Twelfth International Conference on Machine Learning*. Morgan Kaufmann, 1995.

[4] Nicolas Meuleau, Milos Hauskrecht, Kee-Eung Kim, Leonid Peshkin, Leslie Pack Kaelbling, Thomas Dean, and Craig Boutilier. Solving very large weakly coupled Markov decision processes. In *Proceedings of the Fifteenth National Conference on Artificial Intelligence*, Madison, Wisconsin, 1998. AAAI Press.