Approximation Algorithms for Solving
Cost Observable Markov Decision Processes

Valentina Bayer
Department of Computer Science
Oregon State University
Corvallis, OR 97331
bayer@cs.orst.edu

Partial observability is a result of noisy or imperfect sensors that are not able to reveal the real state of the world. Problems that suffer from partial observability have been modelled as Partially Observable Markov Decision Processes (POMDPs). They have been studied by researchers in Operations Research and Artificial Intelligence for the past 30 years. Nevertheless, solving for the optimal solution or for close approximations to the optimal solution is known to be NP-hard. Current algorithms are very expensive and do not scale well. Many applications can be modelled as POMDPs: quality control, autonomous robots, weapon allocation, medical diagnosis.

Designing approximation algorithms to solve problems that have partial observability is the focus of this research. The model we propose (Cost Observable Markov Decision Processes or COMDPs) associates costs with obtaining information about the current state (Bayer 1998). The COMDP's actions are of two kinds: world actions and observation actions. World actions change the state of the world, but return no observation information. Observation actions return observation information, but the state of the world does not change while performing them.

COMDPs are intended to model situations that arise in diagnosis and active vision where there are many observation actions that do not change the world and relatively few world-changing actions. We are particularly interested in problems where there are many alternative sensing actions (including, especially, no sensing at all) and where, if all observation actions are performed, the entire state of the world is observable (but presumably at very great cost). Hence, COMDPs can also be viewed as a form of fully observable Markov Decision Processes (MDPs) where the agent must pay to receive state information (they are "cost-observable").

Any POMDP can be modeled as a COMDP and vice versa. We want to approximately solve COMDPs and determine what POMDP classes these approximations are good for.

There are two fundamental problems that any POMDP approximation algorithm must address: how to act when "lost" and how to avoid getting "lost".

All current POMDP approximation algorithms address only the first of these two problems. In a situation where the agent is uncertain about the state of the world, they choose an action based on a heuristic such as performing the action whose one-step value-of-information is highest.

We are experimenting with an approximation algorithm for COMDPs that addresses both fundamental problems. Given the underlying fully-observable MDP, the algorithm constructs a new MDP whose reward function incorporates the cost of making observations. Then, by computing the value function of this modified MDP, the "costs of getting unlost" (i.e., the costs of observations) are propagated backward along action trajectories. This allows the agent to choose actions to avoid getting lost. A sequence of MDPs is constructed iteratively, each MDP introducing observation costs based on the preceding MDP. If this sequence of MDPs converges, our method will predict when and how much it is necessary to observe at execution time. We are currently comparing the performance of this algorithm (and variations of it) to optimal policies computed by exact POMDP solution methods and to policies computed by policy-space search. We want to extend this off-line approximation algorithm to an on-line version.

Our long term goal is to apply the COMDP approximation algorithms to the problem of active visual perception in real-time problem-solving tasks such as air traffic control. We also want to look at reinforcement learning algorithms for learning the COMDPs.

Acknowledgements

References

Bayer, V. 1998. Approximation Algorithms for Solving Cost Observable Markov Decision Processes. Technical Report 99-50-01, Oregon State University, Dept. of Computer Science.