Massachusetts Institute of Technology
16.412J/6.834J Cognitive Robotics
Advanced Lecture Proposal

Lawrence Bush, Brian Bairstow and Tony Jimenez
{ bushl2, bairstow}@mit.edu, tjimenez@alum.mit.edu

March 16, 2005

Part A:

Title: Planning in POMDPs.

Team Member List:

- Lawrence Bush
- Brian Bairstow
- Tony Jimenez

Topics:

- An introduction to the fundamentals of POMDPs
- A review and a paper representing the state of the art in POMDP research and a pedagogical explanation of the respective algorithm.
- Demonstration

Part B: Covered Paper

In our lecture, we will be covering the following paper:

Georgios Theocharous and Leslie Pack Kaelbling, "Approximate Planning in POMDPS with Macro-Actions," *Advances in Neural Information Processing Systems 16*, Vancouver, 2004 (NIPS-03).

Part C: Abstract

This lecture will cover partially observable Markov decision processes (POMDP), in particular, as they pertain to macro actions. This algorithm is applied to a robot navigation task, which has a temporal component.

POMDPs are used to model intelligent agents in an uncertain environment. The agent observes its environment, develops a belief state (a probabilistic state estimate) and chooses an action to maximize the expected future reward. POMDPs are powerful because all environments are uncertain to varying degrees. Consequently, POMDPs more closely model reality.

Unfortunately, creating a policy is currently intractable because the number of belief states grows with the number of actual states and the number of actions taken. This has kept POMDPs from being utilized in real-world applications. One way to ease the computational complexity is to reduce the size of the state space using macro actions. This is based on the notion that an agent usually only travels through a small part of its belief space.

Macro actions are commands like "move down the hallway" or "move to the site" rather than "move forward two inches" or "turn 30 degrees clockwise." Using macro actions effectively turns the state space from a grid to a graph, reducing the number of possible states and thus the number of belief states.

Integral to this algorithm are the Q-value update method, the POMDP representation, and the grid based belief space coordinate computation. The specifics of these techniques will also be covered.

Part D: Background Publications

N. Roy, G. Gordon and S. Thrun. ``Finding Approximate POMDP solutions Through Belief Compression''. *Journal of Artificial Intelligence Research,* 23: 1-40, 2005. http://web.mit.edu/nickroy/www/papers/jair05.pdf

Supplementing the above paper:

N. Roy, "PhD Thesis: Finding Approximate POMDP Solutions Through Belief Compression," Robotics Institute, Carnegie Mellon University, 2003. http://mapleleaf.csail.mit.edu/~nickroy/thesis/

Stuart Russell and Peter Norvig, "Artificial Intelligence: A Modern Approach (Second Edition)," Prentice Hall, 2002 http://aima.cs.berkeley.edu/

Leslie Pack Kaelbling, Michael L. Littman and Anthony R. Cassandra, ``Planning and Acting in Partially Observable Stochastic Domains,'' *Artificial Intelligence,* Vol. 101, 1998. http://people.csail.mit.edu/people/lpk/papers/aij98-pomdp.pdf

Hiller and Lieberman, "Introduction to Operations Research (Fourth Edition)," Holden-Day, Inc., 1986 http://www.mhhe.com/engcs/industrial/hillier/

Jaakkola, T., Singh, S., and Jordan, M., "Reinforcement Learning Algorithm for Partially Observable Markov Decision Problems," *Advances In Neural Information Processing Systems*, MIT Press, 1995. http://www.eecs.umich.edu/~baveja/Papers/Nips94b.pdf

Georgios Theocharous, Kevin Murphy, and Leslie Pack Kaelbling, "Representing hierarchical POMDPs as DBNs for multi-scale robot localization," *International Conference on Robotics and Automation*, 2004. http://people.csail.mit.edu/people/lpk/papers/theochar-icra04.pdf

Part E: Division of Labor

Primary Activities:

- POMDP Explanation (Tony)
- Paper representing state of the art research (Brian)
- Demonstration (Larry)

Overarching Activities:

- Pedagogical Slide Annotations
- Slide Construction
- Presentation Flow and Cohesiveness
- Presentation Effectiveness

Part F: Demonstration

We intend to create a pedagogical demonstration of POMDPs. This is to take the form of a visualization of the evolving decision process and final action sequence of a POMDP on an instructive (small) path-planning problem. This edification process is intended to instruct the listener while building the plan.

For this demonstration, we intend to use "Perseus," a set of Matlab functions implementing a randomized point-based approximate value iteration algorithm for Partially Observable Markov Decision Processes (POMDPs).

http://staff.science.uva.nl/~mtjspaan/software/approx/