

Study Questions

April 18, 2004

- Name and describe the four components of a Markov Decision Process.
- After each step (each action choice) in an MDP a learning agent receives four pieces of information. What are they?
- What is the objective of a reinforcement learning algorithm?
- What is $Q^*[s, a]$? What does the * mean?
- Suppose we have an MDP with two states s_1, s_2 and two actions a_1, a_2 . We are using Q-learning to figure out how to solve the problem. Currently $Q[s_1, a_1] = 1.0$, $Q[s_1, a_2] = -1$, $Q[s_2, a_1] = -1$ and $Q[s_2, a_2] = 1.0$. Suppose we are in state s_1 , take action a_1 and we end up in s_2 with a reward of 2.0. Which Q-value should change and what should it change to? Assume the learning rate and discount rates are $\alpha = 0.2, \gamma = 0.9$, respectively.
- Suppose you were using reinforcement learning to program a robot to solve a series of maze problems. For each of the maze problems the same maze is used, but a different goal location is selected. Would you use model-based or model-free learning to solve this problem? Why?
- What is Tucker's favorite equation? What general class of problems can it be used to solve?
- We need a way to localize a robot with a sensor that detects known landmarks, but it can only evaluate distance to them. What sort of localization scheme would you use? The robot currently thinks it is at $(3.0, 2.0)$ it sees a landmark it knows is at $(0, 2.0)$. The sensor says the landmark is 2.0 units away. How would your algorithm use this information to update the robot's estimated location? What is the new, updated estimated location?