Reinforcement Learning
Semester 2, 2011/12
Coursework Assignment 2

Your task is to develop a Matlab program to solve a simple navigation problem using the QMDP model and algorithm. The environment has 12 discrete states. Initially, the robot is placed at a random location, chosen uniformly among the possible 12 states. Its goal is to advance to state 7, as shown.

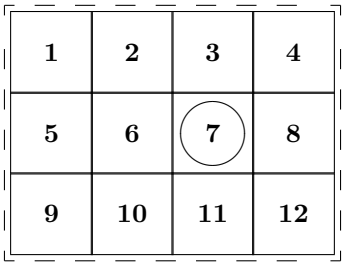

Figure 1: The environment with 12 discrete states. The target is state 7.

At any point in time, the robot may go along the canonical directions - north, east, south or west. Its only sensor is a bumper: when it hits a wall, the bumper triggers and the robot does not change state. Other than this piece of weak information, the robot lacks any ability to sense what state it is in. Also, it cannot sense the direction of its bumper. There is no further noise or uncertainty in this problem - just the initial location uncertainty and weak sensing.

1. To start with, assume that the robot can observe its location. Set up a suitable MDP and find an optimal policy for reaching the target state, using value iteration. Show your value function and the resultant policy, in addition to salient design decisions related to the MDP.

2. Next, with the robot unable to observe its location, implement (in Matlab) a Bayes filter that tracks the robot belief over its location while it moves using a random navigation policy. The filter should use the observation of bumps, the actions taken, and the knowledge of the world map to update the belief over the state.

3. Implement the QMDP algorithm to solve the navigation *PO*MDP. The robot starts in an unknown location, and should advance toward state 7.

An episode only completes if the robot reaches the target *and* its belief has converged. How many steps does it take the robot (on average) to reach the target?

4. Comment on the performance of your algorithm; does it always achieve the specified task? What improvements can you propose to make it do so?

Subramanian Ramamoorthy and Majd Hawasly
9th March 2012