CSE 574 Homework 3

Zackary Crosley

September 19, 2018

Problem 1. Markov Decision Processes: Consider the world shown in the figure. Assume that 80% of the time the agent moves in its intended direction and 10% of the time it moves in each of the two right angles to that direction. Implement value iteration for each r value below. Discount rewards with a discount factor 0.99. Show the policy obtained in each case and explain why that policy was the intuitive output.

1. r = 100

2. r = -3

3. r = 0

4. r = 3

Problem 2. Value Iteration: For the zero-sum, turn-taking game in Homework 2, let R(s) be the reward for A in state s. Let -R(s) be the reward for B with A in state s. Let $U_A(s)$ be the utility of state s when it's A's turn to move in s, and $U_B(s)$ the same for B.

1. Write down Bellman equations defining $U_A(s)$ and $U_B(s)$.

2. Explain how to do two player value iteration with these conditions and define suitable termination criteria.

3. Draw the state space showing moves by A in solid lines and moves by B in dashed lines. Mark each state with R(s). Arrange states sA sB on a grid, using sA and sB as coordinates.

4. Apply two-player value iteration to solve the game, and derive the optimal policy.

Problem 3.

Problem 4.

Problem 5.

Problem 6.

Problem 7.