

CMSC 478/678 Machine Learning - Fall 2014

Homework Assignment 7

Due at the start of class on November 6th

(Reinforcement Learning) Implement Q-learning (with the programming language of your choice) and use it to find an optimal policy for traversing a 15x15 grid world with 4 actions that move one step in the standard compass directions. Each episode starts with the learner in the upper left corner state and ends when the learner enters the lower right corner state. Attempts to move off the edges of the grid world should result in the agent staying in the same state. Choose any reward structure that causes the learner to minimize the number of steps required to end the episode. Turn in the following:

- Your code
- List of the parameters used in learning, such as rewards, discount factor, learning rate, exploration probability, and anything else that seems relevant
- A learning curve that shows the number of steps until the goal state is reached on the vertical axis and the episode number on the horizontal axis