Northeastern University

CS 7180 – Special Topics in AI (Reinforcement Learning)

Fall 2018, Robert Platt

Dynamic Programming Assignment

Name: _____

| Problem	Points
V^*, Q^*, and π^*	/20
Working the numbers for one iteration of dynamic programming.	/10
Programming: Value iteration	/20
Total	/50

Instructions

- Don't cheat!

(20 pts.) V^*, Q^*, AND π^*

a. Give an equation for V^* in terms of Q^*. Break it down into a form with no expectations.

b. Give an equation for Q^* in terms of V^*. Break it down into a form with no expectations.

c. Give an equation for π^* in terms of Q^*, assuming that π^* is deterministic.

d. Give an equation for π^* in terms of V^*, assuming that π^* is deterministic.

(10 pts.) WORKING THE NUMBERS FOR ONE ITERATION OF DYNAMIC PRO-
GRAMMING.

Consider the following fragment of an MDP graph. The fractional numbers indicate the worlds transition probabilities and the whole numbers indicate expected rewards. The three numbers at the bottom indicate what you can take to be the value of the corresponding states. The discount rate is 0.9. Let π be the equiprobable random policy (all actions equally likely). What is the value of the top node for this policy and for the optimal policy? That is, what are $V^\pi(x)$ and $V^*(x)$? Show your work.

$$v_\pi(x) =$$

$$v_*(x) =$$
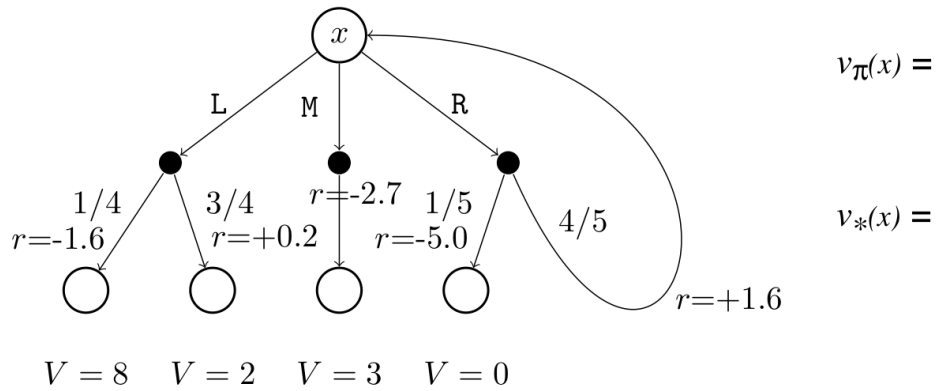
(20 pts.) PROGRAMMING: VALUE ITERATION

Complete the two programming excersises in *cs7180-hw4-dynamic-programming-solution.ipynb*.