

# Dynamic Programming

ECE/CSE RL Monsoon 2019

Sep 3, 2019

**Question 1.** Solve Exercise 3.4. Explain how you obtained the table. Your solution may be hand-written.

**Question 2.** Write code that solves the linear equations required to find  $v_\pi(s)$  and generate the values in the table in Figure 3.2. Note that the policy  $\pi$  picks all valid actions in a state with equal probability. Add comments to your code that explain all your steps.

**Question 3.** Solve Exercises 3.15 and 3.16.

**Question 4.** Write code that generates the optimal state-value function and the optimal policy for the Gridworld in Figure 3.5. You want to solve the corresponding system of non-linear equations. Explain all your steps.

**Question 5.** Given an equation for  $v_*$  in terms of  $q_*$ .

**Question 6.** Code policy iteration and value iteration (VI) to solve the Gridworld in Example 4.1. Your code must log output of each iteration. Pick up a few sample iterations to show policy evaluation and improvement at work. Similarly, show using a few obtained iterations that every iteration of VI improves the value function. Your code must include the fix to the bug mentioned in Exercise 4.4.

**Question 7.** Code exercise 4.7.