## CPSC 4820/6820 Intro to Al

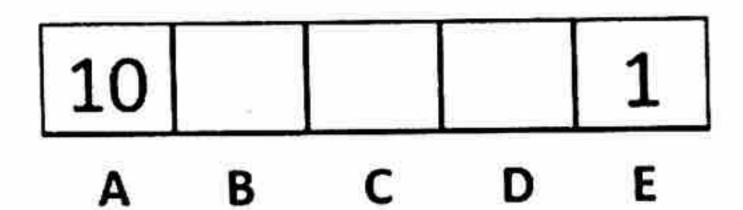
Assigned: Feb 15, 2019

Due: Feb 21, 2019 (start of class)

Quiz 4

(Grading: 0-10 points)

## 1. MDPs



You are given the 1D grid world shown above where you can go east, west, or exit. You can only exit at states A and E receiving a reward of 10 and 1, respectively. Assuming that the world is deterministic (every action you take will always succeed), and the living reward in non exit states is 0, please answer the following questions:

(i) Given a discount factor  $\gamma = 1$ , what is the optimal policy?

(ii) Given a discount factor  $\gamma = 0.1$ , what is the optimal policy?

Г		ACTION			
1	10	+	4	1	1
		250			-

$$A = \emptyset + \cdot 1(10) = 1$$
  
 $B = \emptyset + \cdot 1(0) + \cdot 01(10) = \cdot$ 

## 2. Value Iteration

Consider the same setup as above, though the world now is stochastic. When the agent moves east or west, it has 90% chance of succeeding and 10% of staying where it is. For example, if the agent takes the west action from state B, there is a 90% chance to land on state A, and 10% chance to stay put. Assuming no discounts ( $\gamma = 1$ ), fill in the following table of value iteration values for the first 4 iterations.

States	A	В	С	D	E
$V_{0}$	Ø	Ø	Ø	Ø	Ø
V <sub>1</sub>	+ \O	Ø	Ø	Ø	+ 1
$V_2$	+10	9	Ø	• 9	+ 1
$V_3$	+10	9.9	8.1	. 99	+ 1

$$V_{\mathbb{Z}} = (-9)(10) + (-1) V_{\mathbb{Z}}(B)$$

$$C = (-9) V(B) + (-1) V_{\mathbb{Z}}(B)$$

$$= (-9) V(B) + (-1) V_{\mathbb{Z}}(B)$$

$$C = (.9)(10) + (.1) V_{2}(B)$$

$$C = (.9) V(B) + (.1) V(C)$$

$$D = (.9) V(B) + (.1) V(C)$$

$$Q(9)(10) + .1(9)$$

$$Q(9)(10) + .1(9)$$

$$Q(1) + .1(9)$$

$$Q(1) + .1(9)$$

$$Q(1) + .1(9)$$