Natalie DellaMaria

CPSC 4820/6820
Intro to AI

Assigned: Feb 15, 2019
Due: Feb 21, 2019 (start of class)

Quiz 4

(Grading: 0–10 points)

1. MDPs

10				1
A	B	C	D	E

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?

10	←	←	←	1

A = ∅ + 1(10) = 10
B = ∅ + 1(∅) + 1(10) = 10
C = ∅ + 1(∅) + 1(∅) + 1(10) = 10

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

10	←	←	→	1

A = ∅ + .1(10) = 1
B = ∅ + .1(0) + .01(10) = .1 || 0 + .1(0) + .01(1) = .01
C = ∅ + .1(0) + .01(0) + .001(10) = .01 || 0 + .1(1) = .1

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	B	C	D	E
V_0	∅	∅	∅	∅	∅
V_1	+10	∅	∅	∅	+1
V_2	+10	9	∅	.9	+1
V_3	+10	9.9	8.1	.99	+1

.1
.9
.09

$V_2(B) = (.9)(10) + (.1)V_2(B)$ (.9)(10) + .1(9)
$C = (.9)V(B) + (.1)V(C)$ .9(9) + .1(0) || .9(.9) + .1(0)
$D = (.9)V(B) + (.1)V(C)$ .9(1) + .1(.9)
.9 + .09