

Thursday, April 24, 2014
CS561 Artificial Intelligence
Quiz 9b

Name:
Student ID:
Student USC email:

This study resource was
shared via CourseHero.com

Given the grid below answer the questions below:

3	0.812	0.868	0.918	<div>+ 1</div>
2	0.762		0.660	<div>- 1</div>
1	0.705	0.655	0.611	0.388
	1	2	3	4

1) (5pts) What would be the optimal policy for the grid?

1,1 N 2,1 W 3,1 N 4,1 W

1,2 N 3,2 N

1,3 E 2,3 E 3,3 E

2) (5pts) What are the Q-values for the square (3,2) given that the agent is deterministic, i.e. $Pr(s, a, s') = 1$, $R(3,2)=0$ and $\gamma = .5$

Recall that $Q(a, s) = \sum Pr(s, a, s')[R(s) + \gamma \max Q(a', s')]$ and $Q(a, s) = 0$ at Terminal states (4,2), (4,3)

$$Q(N, (3,2)) = 1 * [0 + .5 (Q(E, (3,3)))] = .25$$

$$Q(E, (3,2)) = 1 * [-1 + .5 (1)] = -.5$$

$$Q(S, (3,2)) = 1 * [0 + .5 (Q(N, (3,1)))] = .0625$$

$$Q(E, (3,3)) = 1 * [0 + .5 (1)] = .5$$

$$Q(N, (3,1)) = 1 * [0 + .5 (Q(N, (3,2)))] = .125$$