Artificial Intelligence

CS 470

Instructor – Michael A. Goodrich

Name:_____

1. (*10 points*) (*45 minutes*) Do problem 17.1 in the book. Ignore the part about "projecting a hidden Markov model.".

2. (*10 points*) (*60 minutes*) Do problem 17.4 in the book. For part b, rather than do the problem for an undiscounted MDP, assume $\gamma = 0.9$. Experience suggests that using a lot of paper will save you a lot of time.

3. (*10 points*) (*45 minutes*) Consider the following world.

13	14	15	16
9	10	11	12
5	6	7	8
1	2	3	4

In the figure, the numbers in each cell represent states. For example, the cell in the lower left hand corner is state s_1. The goal is in state s_{13} at the upper left. Consider the following policy:

$$\pi = \begin{cases} N & \text{for states 1-12} \\ W & \text{for states 13-16} \end{cases}.$$

The policy does not need to specify what to do in s_{13} since we assume that this state is an "absorbing state", meaning that once this state is entered the agent will never leave. The reward structure for this problem is defined as follows:

$$r(s,a) = \begin{cases} 1 & \text{when state is } s_{14} \text{ and action is } W \\ 1 & \text{when state is } s_9 \text{ and action is } N \\ 0 & \text{otherwise} \end{cases}.$$

The world is deterministic except at state s_1 and state s_6. At these states, the transition probability is as follows.

$P(s'\|s_1, N)$	$P(s_5\|s_1, N) = 0.8$	$P(s_2\|s_1, N) = 0.2$
$P(s'\|s_6, N)$	$P(s_{10}\|s_6, N) = 0.7$	$P(s_7\|s_6, N) = 0.3$

Find the expected value, $U(s, \pi)$, of applying this policy from state s_1 given a discount parameter of $\gamma = 0.9$.

4. (*10 points*) (*10 minutes*) Which actions are not dominated?

Utility/action	a_1	a_2	a_3	a_4	a_5
$\mathcal{U}_1$	1	2	3	4	5
$\mathcal{U}_2$	-1	-2	-3	-4	-5
$\mathcal{U}_3$	0	6	0	4	2