Name: **Name:**
Student ID: **Student ID:**

CSE 473 Autumn 2019 HW2

11/27/2019

100 points

Instructions:

1) The homework can be done individually or in a group of two people. Mention the name of your partner in the submission, and you can both submit the same file.
2) We highly recommend typing your homework, but writing and scanning also works.
3) Keep your answers brief but provide enough explanations and details to let us know that you have understood the topic.
4) The assignment is due on November 27.

Topics:	Points
Reinforcement Learning I	25
Reinforcement Learning II	15
Uncertainty I	20
Uncertainty II	15
Bayesian Modeling	25

Problem 1. Card Game Q-Learning [25 points]

You are playing a peculiar card game, but unfortunately you were not paying attention when the rules were described. You did manage to pick up that for each round you will be holding one of three possible cards [Ace, King, Jack] ([A, K, J], for short) and you can either Bet or Pass, in which case the dealer will reward you points and possibly switch out your card. You decide to use Q-Learning to learn to play this game, in particular you model this game as an MDP with states [A, K, J], actions [Bet, Pass] and discount $\gamma = 1$. To learn the game you use $\alpha = 0.25$.

A) Say you observe the following rounds of play (in order):

s	a	s'	r
A	Bet	K	4
J	Pass	A	0
K	Pass	A	-4
K	Bet	J	-12
J	Bet	A	4
A	Bet	A	-4

What are the estimates for the following Q-values as obtained by Q-learning? All Q-values are initialized to 0. (15 points)

i) Q(J, Pass) = _____

ii) Q(J, Bet) = _____

B) For this next part, we will switch to a feature based representation. We will use two features:

$$f_1(s, a) = 1$$

$$f_2(s, a) = \begin{cases} 1 & a = \text{Bet} \\ 0 & a = \text{Pass} \end{cases}$$

Starting from initial weights of 0, compute the updated weights after observing the following samples:

s	a	s'	r
A	Bet	K	8
K	Pass	A	0

What are the weights after the first update, in other words, after using the first sample? (5 points)

i) w1 = _____

ii) w2 = _____

What are the weights after the second update, in other words, after using the second sample? (5 points)

iii) w1 = _____

iv) w2 = _____

Problem 2. Reinforcement Learning [15 points]

A) For each of the following statements, please specify if the statement is true or false and then explain briefly why you chose that answer. (8 points)

i) [true or false] A large discount (γ close to 0) encourages greedy behavior.

ii) [true or false] In a deterministic MDP (i.e. one in which each state / action leads to a single deterministic next state), the Q-learning update with a learning rate of α = 1 will correctly learn the optimal q-values (assume that all state/action pairs are visited sufficiently often)

iii) [true or false] A discount < 1 can always be expressed as a negative living reward.

iv) [true or false] Q-learning is an off-policy learning algorithm, and therefore converges only when the agent begins to act optimally.

B) Given the following list of Q-values for state s and the set of actions {*Left*, *Right*, *Fire*} (7 points):

Q(s, *Left*) = 0.15
Q(s, *Right*) = 0.95
Q(s, *Fire*) = 0.5

What is the probability that we will take each action on our next move when following an ε-greedy exploration policy (assuming all random movements are chosen uniformly from all actions)?

Action	Probability, in terms of ε
Left	
Right	
Fire	

Problem 3. Uncertainty - Conditional Variables [20 points]

For these problems, assume we have three random variables A, B, C with possible instantiations a, b, c, respectively.

1. The conditionalized version of the general product rule is P(a,b|c) = P(a|b,c)P(b|c). Show how to derive this rule using the definition of conditional probability. (5 points)

2. If P(a,b,c) = 0.01, P(a|b,c) = 0.2, and P(b|c) = 0.1. what is P(c)? (5 points)

3. If P(c) = 0.3, P(b) = 0.15 , P(b|c) = 0.1, and P(a|b,c) = 0.2 what is P(a,b|c)? (5 points)

4. Prove that the two definitions of conditional independence are equivalent (5 points) i.e., that P(a, b|c) = P(a|c) P(b|c) is equivalent to P(a|b,c) = P(a|c) for all a, b, c.

Problem 4. Uncertainty - UW Crows [15 points]

Assume 65% of all UW campus birds are crows. Additionally, assume that 15% of these crows have an ankle band and only 10% of non-crow campus birds have an ankle band.

A) What is the probability that a bird is a crow and has an ankle band? (5 points)

B) What is the probability that a bird is a crow if it is known to have an ankle band? (5 points)

C) What is the probability that a bird is unbanded or is a crow? (5 points)

Problem 5. Bayesian Modeling [25 points]

Consider the following modeling challenge, which might arise when working on a homework problem. You can either study hard or trick the TA into telling you the right answer. Given these choices, you might actually learn the material and you might also get a passing grade on the problem. Remember, you won't learn materials if you just trick the TA. All of these possibilities will happen with certain probabilities, that you are welcome to imagine however you like.

A) Model this problem as a Bayesian network representing a joint distribution over four binary random variables. Since there could be more than one possible answer, briefly motivate your choices. (15 points)

B) Write three independence assumptions, or conditional independence assumptions, that your network encodes. (10 points)