Name: **Student ID:**

CSE 573 Winter 2021 HW2

2/25/2021

100 points

Instructions:

1) The homework should be done individually. Don't forget to write your name.
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 Friday, March 12.

Topics:	Points
Value Iteration	12
Reinforcement Learning	16+9
Uncertainty	22
Bayesian Modeling and Inference	12
HMMs	29 (4)*
Extra credit: Bayesian Inference	(10)*

* Extra credit

Value Iteration [12 Points]

Consider the 101x3 world below. In the start state, the agent has a choice of two deterministic actions, *Up* or *Down*, but in the other states the agent always takes the deterministic action, *Right*.

-50	1	1	1	1	..	1	1	1	1(**TERMINAL**)
START									
50	-1	-1	-1	-1	..	-1	-1	-1	-1(**TERMINAL**)

1. (6 pts) Compute the utility of each action as a function of γ.
2. (6 pts) Assuming a discounted reward function, for what values of the discount factor γ should the agent choose Up as the initial action?

Reinforcement Learning [16 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

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

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

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

i) w1 = ________________________
ii) w2 = ________________________

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

iii) w1 = ________________________
iv) w2 = ________________________

Reinforcement Learning [9 pts]

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	

Probability and Uncertainty [22 points]:

A. (4 pts) Suppose Boolean random variables A and B are independent of each other (They can only have two values of True and False). Determine the missing entries x and y in the joint distribution of P(A, B) shown below.

$$P(A = T,\ B = T) = 0.4$$
$$P(A = T,\ B = F) = 0.1$$
$$P(A = F,\ B = T) = x$$
$$P(A = F,\ B = F) = y$$

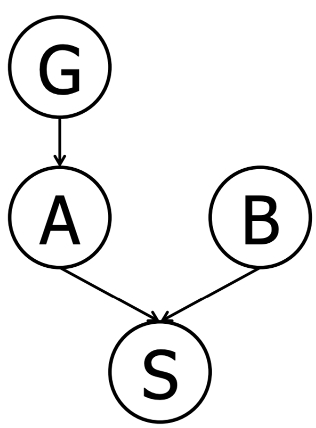
B. For these problems, assume we have three random variables A, B, C with possible instantiations a, b, c, respectively.
   1. (5 pts) The conditionalized version of the general product rule is P(a,b|c) = P(a|b,c)P(b|c). Show how to derive this rule using the definition of conditional probability.
   2. (4 pts) If P(a,b,c) = 0.01, P(a|b,c) = 0.2, and P(b|c) = 0.1. What is P(c)?
   3. (5 pts) Prove that the two definitions of conditional independence are equivalent 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.
   4. (4 pts) Expand P(A, B, C) given that $A \perp B | C$ and $A \perp C$.

Bayes' Nets Representation [12 Points]

Suppose that a patient can have a symptom (S) that can be caused by two different diseases (A and B). It is known that the variation of gene G plays a big role in the manifestation of disease A. The Bayes' Net and corresponding conditional probability tables for this situation are shown below. For each part, you may leave your answer as an arithmetic expression.

$\mathbb{P}(G)$	
$+g$	0.1
$-g$	0.9

$\mathbb{P}(A\|G)$		
$+g$	$+a$	1.0
$+g$	$-a$	0.0
$-g$	$+a$	0.1
$-g$	$-a$	0.9



$\mathbb{P}(B)$	
$+b$	0.4
$-b$	0.6

$\mathbb{P}(S\|A,B)$			
$+a$	$+b$	$+s$	1.0
$+a$	$+b$	$-s$	0.0
$+a$	$-b$	$+s$	0.9
$+a$	$-b$	$-s$	0.1
$-a$	$+b$	$+s$	0.8
$-a$	$+b$	$-s$	0.2
$-a$	$-b$	$+s$	0.1
$-a$	$-b$	$-s$	0.9

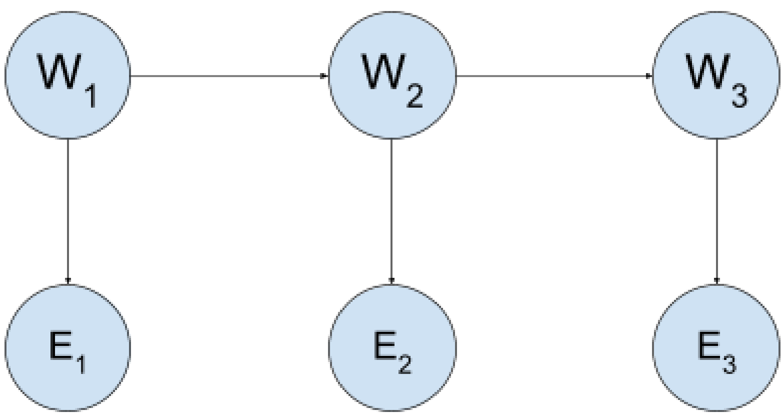
a) (3pts) Compute the following entry from the joint distribution: P(+g, +a, +b, +s)=?
b) (3pts) What is the probability that a patient has disease A? P(+a) = ?
c) (3pts) What is the probability that a patient has disease A given that they have disease B? P(+a|+b) = ?
d) (3pts) What is the probability that a patient has disease A given that they have symptom S and disease B? P(+a |+s, +b) = ?

Hidden Markov Models [17 Points]

Consider the following HMM, where W_1, W_2 and W_3 are state variables, and E_1, E_2 and E_3 are evidence variables.

W_1	$P(W_1)$
0	0.2
1	0.8

W_t	W_{t+1}	$P(W_{t+1} \| W_t)$
0	0	0.9
0	1	0.1
1	0	0.7
1	1	0.3



W_t	E_t	$P(E_t \| W_t)$
0	A	0.6
0	B	0.4
1	A	0.3
1	B	0.7

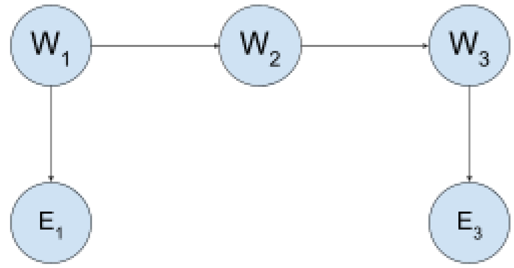
Using probabilities given in the tables,

A. (3pts) Compute $\mathbf{a^i} = P(W_1 = i \mid E_1 = A)$ for i=0 and i=1.

B. (3pts) Compute $\mathbf{b^i} = P(W_2 = i \mid E_1 = A)$ for i=0 and i=1; you may express your solution in terms of $\mathbf{a^0}$ and $\mathbf{a^1}$.

C. (3pts) Compute $\mathbf{c^i} = P(W_2 = i \mid E_1 = A, E_2 = B)$ for i=0 and i=1; you may express your solution in terms of $\mathbf{a^0}$, $\mathbf{a^1}$, $\mathbf{b^0}$, $\mathbf{b^1}$.

Now consider this modified HMM, where we do not receive evidence at time 2.

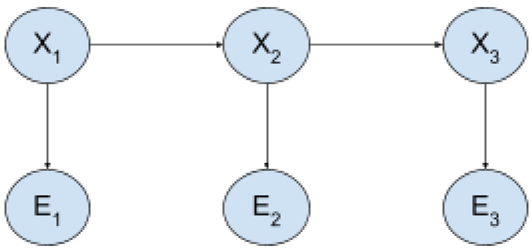


D. (4pts) Compute **d**^{i}=P(W_3 =i| E_1=A) for i=0 and i=1, you may express your answers in terms of **a**^{i} and **b**^{i}.

E. (4pts) Compute **e**^{i}=P(W_3 =i| E_1=A, E_3=B) for i=0 and i=1, you may express your answers in terms of **a**^{i}, **b**^{i}, and **d**^{i}.

Modified HMM (16 Points)

A standard First-Order HMM is graphically represented like this, where the hidden variable at time t only depends on the hidden variable at the previous time step, and every state only includes one evidence variable.



The forward algorithm for a standard HMM is shown below:

$$P(X_t|e_{1:t-1}) = \sum_{x_{t-1}} P(X_t|x_{t-1})P(x_{t-1}|e_{1:t-1})$$

$$P(X_t|e_{1:t}) \propto P(X_t|e_{1:t-1})P(e_t|x_t)$$

Where the first equation is the Elapsed Time and the second equation is the Observation update.

Now imagine a modified HMM, where we receive two independent pieces of evidence each time step. Let A and B be random variables representing those pieces of evidence at time t.

A. (4pts) Draw three steps of a modified HMM chain for this settings. Your drawing should include the variables X, A, and B:
B. (4pts) Write down the formula for Observation for this setting:

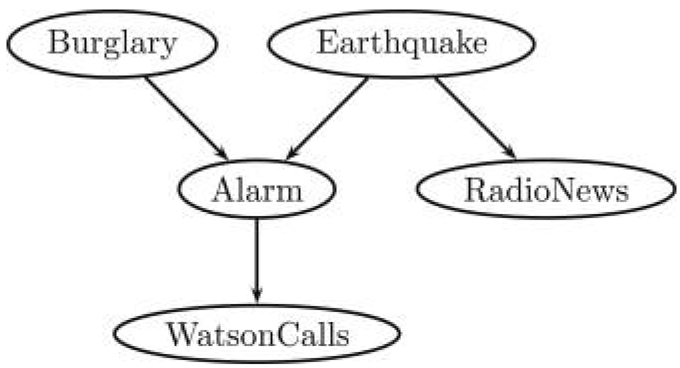
$P(X_t|a_{1:t}, b_{1:t}) \propto$

Now consider a different setting where there is only one evidence variable, E, but the next state depends on the evidence from our previous time step as well as the previous hidden state. You could imagine a robot that changes its behavior based on what its sensors are reporting.

C. (4pts) Draw three steps of a modified HMM chain for this setting. Your drawing should include the variables X and E.
D. (**Extra Credit:** 4pts) Write down the formula for the Elapsed Time update for this setting: $P(X_t|e_{1:t-1}) =$

Extra Credit: Bayesian Inference [10pts]

Consider the following Bayesian Network, with variables B, E, A, R, and W as a modified version of the Alarm network explained in class.



Consider computing the query P(B|W=true)

A) (3pts) Write the set of initial factors that would be created, after incorporating evidence. You do not need to write the full tables of numbers for each factor, just clearly indicate the function signature, e.g. P(X,Y,Z).
B) (4pts) Write, in order, the signatures for the new factors that get created when running the variable elimination algorithm with the variable elimination ordering A, E, R.
C) (3pts) Provide, if possible, a variable ordering that is more computationally efficient (as measured by total # of rows in factor tables) than the one in part B.