University of Regina

Sample Final Examination

ENSE496AC Artificial Intelligence

December 2019 Time Allowed: 2 hours

INSTRUCTIONS:

1. This paper consists of **8** questions and **11** pages.
2. Answer **ALL** questions.
3. Write your answers in the space provided on the question paper. If there is insufficient space, you may use the additional writing space on page 12.
4. This is an OPEN BOOK exam. You are allowed to bring in ONE electronic device (e.g. laptop, notebook, tablet, mobile phone, etc.) but you must DISABLE all wired or wireless connectivity (e.g. cellular, WiFi, Bluetooth, NFC, etc.). You are also NOT allowed to draw or write any code on your device, NOR use the camera function on your device.
5. Write your name and student ID here:

Name : ____

Student ID : ____

1. (20 marks) :
2. (10 marks) :
3. (10 marks) :
4. (10 marks) :
5. (10 marks) :
6. (10 marks) :
7. (10 marks) :
8. (20 marks) :
Total (100 marks) :

1. (a) Consider a modified version of the racing car example (Fig. Q1) where:

- There are 3 States : “Cool”, “Warm” and “Overheated”;
- There are 2 Actions: “Slow” and “Fast”;
- When a “Cool” car goes “Slow”, it has a 100% chance of remaining in the “Cool” state with a reward of +1;
- When a “Cool” car goes “Fast”, it has a 50% chance of remaining in the “Cool” state with a reward of +2, and a 50% chance of changing into the “Warm” state with a reward of also +2;
- When a “Warm” car goes “Slow”, it has a 50% chance of remaining in the “Warm” state with a reward of +1, and a 50% chance of changing into the “Cool” state with a reward of also +1;
- When a “Warm” car goes “Fast”, it has a 50% chance of remaining in the “Warm” state with a reward of +2, and a 50% chance of changing into the “Overheated” state with a reward of –10;
- When a car is in the “Overheated” state, it goes out of action and it is irreversible;

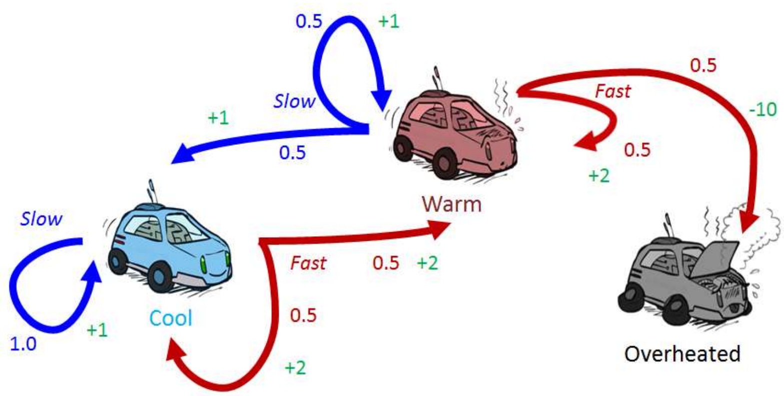


Fig. Q1

Use the Bellman equation, or otherwise, find three iterations of the value at each state and fill in the table below. The first iteration, $V_0(s)$, had been filled in for you. Assume that the discount factor, $\gamma = 1$.

V_2			
V_1			
V_0	0	0	0

(10 marks)

(b) Suppose that we have changed a new car and now we do not know the transition probabilities nor the rewards. By test driving the new car several times, we observed the following outcomes and rewards:

Episode 1

s	*a*	*s'*	*r*
Cool	Slow	Cool	1
Cool	Slow	Cool	1
Cool	Fast	Cool	2
Cool	Fast	Cool	2
Cool	Fast	Warm	3
Warm	Fast	Warm	4
Warm	Fast	Warm	4
Warm	Slow	Warm	3
Warm	Slow	Warm	3
Warm	Slow	Cool	2

Episode 2

s	*a*	*s'*	*r*
Cool	Slow	Cool	1
Cool	Slow	Cool	1
Cool	Fast	Cool	2
Cool	Fast	Warm	3
Warm	Fast	Warm	4
Warm	Fast	Warm	4
Warm	Slow	Warm	3
Warm	Slow	Warm	3
Warm	Slow	Cool	2
Cool	Slow	Cool	1

Episode 3

s	*a*	*s'*	*r*
Cool	Slow	Cool	1
Cool	Fast	Cool	2
Cool	Fast	Warm	3
Warm	Fast	Warm	4
Warm	Fast	Warm	4
Warm	Fast	Over heated	-10

Use model-based reinforcement learning, learn the transition probabilities and the rewards, and fill in the table below. Part of the table has been filled in for you. Assume that the discount factor, $\gamma = 1$.

s	*a*	*s'*	$\hat{T}(s,a,s')$	$\hat{R}(s,a,s')$
Cool	Slow	Cool	1.0	
Cool	Fast	Cool		
Cool	Fast			
Warm				
Warm				
Warm				
Warm				

(10 marks)

2. (a) Circle **all** expressions that are equal to **P(A, B, C)**, **given no independence assumptions**:

    (i) P(A | B, C) P(B | C) P(C)

    (ii) P (C | A, B) P(A) P (B)

    (iii) P (A, B | C) P (C)

    (iv) P (C | A, B) P (A, B)

    (v) P (A | B) P (B | C) P (C)

(b) Circle **all** expressions that are equal to **P(A | B)**, **given that A ⫫ B | C:**

    (i) $\frac{P(A \mid C)\, P(B \mid C)}{P(B)}$

    (ii) $\frac{P(A \mid C)\, P(B \mid C)}{P(B|C)}$

    (iii) $\frac{P(A \mid B,C)}{P(A|C)}$

    (iv) $\frac{\sum_c P(B \mid A,C=c)P(A \mid C=c)}{P(B)}$

    (v) $\frac{\sum_c P(A,C=c)P(B \mid C=c)}{\sum_{c\prime} P(A,B,C=c\prime)}$

(10 marks)

3. Suppose we have a Hidden Markov Model (HMM) of 3 states as shown in Fig. Q2, where each state X_i can be either of the letters 'b' or 'e' with equal probability.

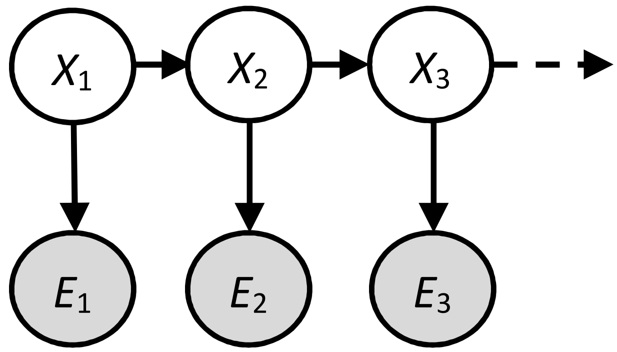


Fig. Q2

The transition probabilities and sensor readings are given as follows:

X_i	X_{i+1}	$P(X_{i+1} \mid X_i)$
b	b	0.4
b	e	0.6
e	b	0.3
e	e	0.7

X_i	E_i	$P(E_i \mid X_i)$
b	b	0.8
b	e	0.2
e	b	0.4
e	e	0.6

Suppose we have observed that the first letter is 'b' and the second letter is 'e'. What is the probability that the third letter is 'e'? Write your answer in the box below.

$P(X_3 = \text{'e'} \mid e_1 = \text{'b'}, e_2 = \text{'e'}) =$ $\boxed{}$

(10 marks)

4. Based only on the structure of the Bayes' Net given in Fig. Q3, circle whether the following conditional independence assertions are guaranteed to be true, guaranteed to be false, or cannot be determined by the structure alone.

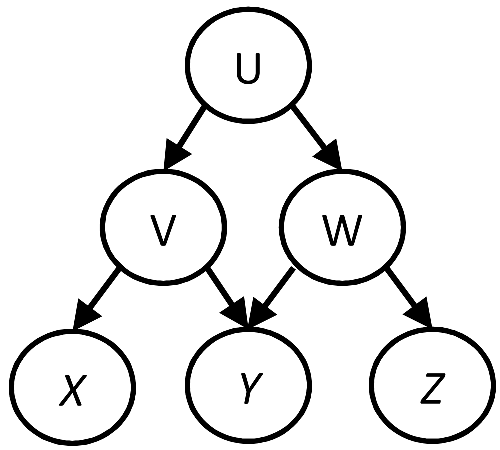
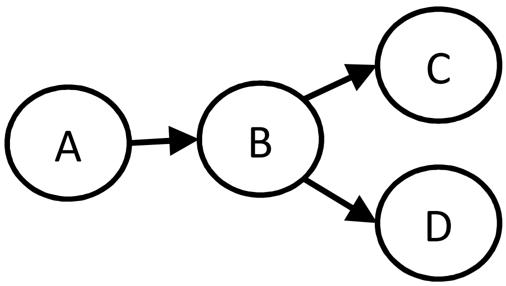


Fig. Q3

(a)	$V \perp\!\!\!\perp W$	Guaranteed true	Guaranteed false	Cannot be determined
(b)	$V \perp\!\!\!\perp W \mid U$	Guaranteed true	Guaranteed false	Cannot be determined
(c)	$V \perp\!\!\!\perp W \mid Y$	Guaranteed true	Guaranteed false	Cannot be determined
(d)	$U \perp\!\!\!\perp Y$	Guaranteed true	Guaranteed false	Cannot be determined
(e)	$U \perp\!\!\!\perp Y \mid V$	Guaranteed true	Guaranteed false	Cannot be determined

(10 marks)

5. Assume the following Bayes' net and the corresponding distributions over the variables in the Bayes' net in Fig. Q4:



A	P(A)
+a	1/5
-a	4/5

A	B	P(B|A)
+a	+b	1/5
+a	-b	4/5
-a	+b	1/2
-a	-b	1/2

B	C	P(C|B)
+b	+c	1/4
+b	-c	3/4
-b	+c	2/5
-b	-c	3/5

B	D	P(D|B)
+b	+d	1/2
+b	-d	1/2
-b	+d	4/5
-b	-d	1/5

Fig. Q4

(a) Your task is now to estimate P(+a | -b, -c, -d) using rejection sampling. Below are some samples that have been produced by prior sampling (that is, the rejection stage in rejection sampling hasn't happened yet). Cross out the samples that would be rejected by rejection sampling (the first one has been done for you):

~~– a – b + c + d~~   – a – b – c – d   + a – b – c + d

– a + b – c – d   – a – b + c – d   + a – b – c – d

(b) Using the above samples, what value would you estimate for P(+a | -b, -c, -d) using rejection sampling?

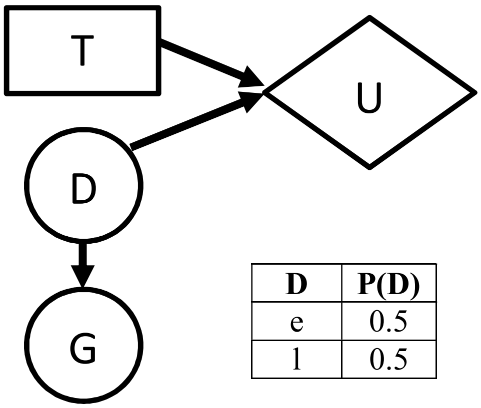
(c) Using the following samples (which were generated using likelihood weighting), estimate P(+a | -b, -c, -d) using likelihood weighting, or state why it cannot be computed.

+ a – b – c – d
– a – b – c – d
+ a – b – c – d

P(+a | -b, -c, -d) =

(10 marks)

6. You want to catch a bus to Downtown this weekend but the bus departure time (D) might be early (e) or late (l). You must decide on the time (T) to start waiting at the bus station, which is also either early (e) or late (l). The bus guide (G) gives you some departure time information, but it is not accurate. The probability values and utilities are shown in Fig. Q5.



D	P(D)
e	0.5
l	0.5

D	G	P(G|D)
e	e	0.8
e	l	0.2
l	e	0.4
l	l	0.6

D	T	U(D,T)
e	e	600
e	l	0
l	e	300
l	l	600

G	P(G)
e	0.6
l	0.4

G	D	P(D|G)
e	e	2/3
e	l	1/3
l	e	1/4
l	l	3/4

Fig. Q5

Two additional tables, P(G) and P(D|G), have been computed for you. You may find them useful to answer the following questions. Compute the following quantities:

(a) First, we consider the case that you do not have the bus guide. Compute:

(i) EU(T = e) =

(ii) EU(T = l) =

(iii) MEU({}) =

(b) Next, we consider the case where you decide to check the bus guide. Compute:

(i) EU(T = e | G = e) =

(ii) EU(T = l | G = e) =

(iii) MEU({G = e}) =

(iv) EU(T = e | G = l) =

(iv) EU(T = l | G = l) =

(v) MEU({G = l}) =

(c) What is the value of the bus guide's information?

VPI(G) = (10 marks)

7. Consider a multi-class perceptron with initial weight vectors $w_A = (0, 0, 0)$, $w_B = (0, 0, 0)$, and $w_C = (1, 0, 0)$. A training sample is observed, which has a feature vector $f(x) = (1, 2, 0)$ and label $y^* = A$.

   (a) What would be the perceptron weight vectors after having seen this training sample?

   $w_A =$

   $w_B =$

   $w_C =$

   (b) Suppose that after a series of training samples, the multi-class perceptron has weight vectors $w_A = (1, 2, 3)$, $w_B = (-1, 0, 2)$, and $w_C = (0, -2, 1)$. A new training example is considered, which has a feature vector $f(x) = (1, -3, 1)$ and label $y^* = B$.

   (i) Which class y would be predicted by the current weight vectors?

   (ii) Would the perceptron update the weight vectors after having seen this training example? If yes, write the resulting weight vectors below:

   $w_A =$

   $w_B =$

   $w_C =$

(10 marks)

8. Suppose a patient can have a symptom (S) that can be caused by two different diseases (A and B). It is known that the presence of gene G directly affects the manifestation of disease A, with 100% certainty. However, disease A could also be caused by other unknown factors.

   (a) Construct the Bayes' Net by drawing the arrows in Fig. Q7:

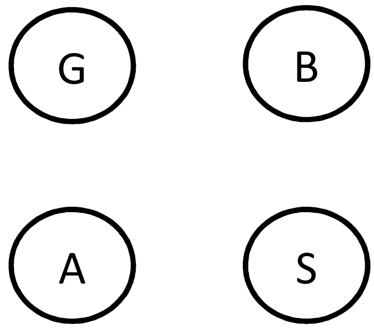


Fig. Q7

   (b) Suppose it is known that the prior probability of a person having the gene G is 0.1, and the prior probabilities of a person having disease A and disease B is 0.19 and 0.4 respectively. Fill in all the entries of the following Conditional Probability Tables:

P(G)	
+g	
-g	

P(B)	
+b	
-b	

P(A|G)		
+g	+a	
+g	-a	
-g	+a	0.1
-g	-a	

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

   (c) What is the probability that a patient has disease A given that they have symptom S?

   (d) What is the probability that a patient has the gene G given that they have symptom S and disease A?

   (e) Consider the query P(G|S),

      (i) What is the most desirable variable elimination ordering? Why?

(ii) In the table below fill in the factors generated at each step of your variable elimination (an example row for eliminating variable G is shown - note that you cannot eliminate variable G because that is the query variable).

	Variable Eliminated	Factor Generated	Current Factors
	None yet	None yet	P(G), P(B), P(A|G), P(S|A, B)
Example	**G**	**f_0(A)**	**P(B), P(S|A,B), f_0(A)**

(iii) Fill in the following Conditional Probability Tables for each of the factors generated:

Example

f_0(A)	
+a	0.19
-a	0.81

f_1		

f_2		

(20 marks)

End of Paper

Additional Writing Space