INDIAN INSTITUTE OF TECHNOLOGY GUWAHATI
Mid-Semester Examination [Jan – May 2024]
CS 561 ARTIFICIAL INTELLIGENCE

Duration: 2hrs **Max Marks: 50**

*Note: Answer **ALL** the questions. Any missing or misprinted data may be assumed suitably. Clearly specify the assumptions while answering. For **PART A**, you must write the **answers** in the appropriate space provided in the **Question paper**. Please avoid overwriting while answering otherwise your marks will be deducted. For **PART B**, use the blank answer booklet. Return **BOTH** question paper and answer script together to the invigilator.*

Name:	**Roll No.**
Student's Signature:	**Invigilator's Signature:**

PART A

1. For each of the following questions tick all the correct answers:

(i) The environment for the following task(s) is not dynamic **[2]**

Crossword Puzzle	☐	Refinery controller	☐
Part-picking robot	☐	Tic-tac-toe game	☐

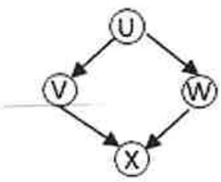
(ii) Which of the following statement(s) is correct? **[2]**

☐ Adding an edge to a Bayes net will always strictly increase the number of distributions the Bayes net can represent.

☐ Every Bayes net with the same number of edges has the same number of independences.

☐ Each variable in a Bayes net is conditionally independent of its non-descendants, given its parents.

(iii) Which of the following conditional independence(s) are asserted by the following Bayes net? **[2]**



☐ V is conditionally independent of W given U

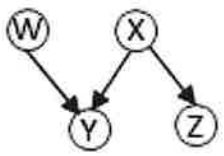
☐ X is conditionally independent of U given V and W

PART B

2. Because likelihood weighting uses all the samples generated, it can be much more efficient than rejection sampling. It will, however, suffer a degradation in performance as the number of evidence variables increases. Do you agree with this? Justify your answer. **[4]**

3. Show that the target distribution p(x) is the stationary distribution of the Markov chain defined by the Metropolis-Hastings algorithm by showing that the detailed balance is satisfied. **[6]**

4. Consider the following Bayes net with random variable W, X, Y, Z. [2+3+2+3] **[10]**



P(W) = 0.5	
P(X) = 0.5	

X	P(Z|X)
0	0.4
1	0.6

W	X	Y	P(Y|W,X)
0	0	0	0.6
0	0	1	0.4
0	1	0	0.4
0	1	1	0.6
1	0	0	0.8
1	0	1	?
1	1	0	0.4
1	1	1	0.6

(i) Determine P (Y = 1|W = 1, X = 0) ('?' entry in the Table) and P (W = 1|X = 0, Y = 1).
(ii) Consider rejection sampling, give a valid and most efficient topological order to estimate P (Y = 1|Z = 1). Justify your answer.
(iii) What will be the weight of the sample (W = 0, X = 0, Y = 0, Z = 1), if you consider likelihood weighting for sampling for the same query as in (ii).
(iv) Consider Gibbs sampling for the same query as in (ii), we initialize W = 0, X = 0, Y = 0, Z = 1, and choose to re-sample W. What is the probability that we still get W = 0 after re-sampling.

5. Consider the Bayesian Network given in Figure (i). Use variable eleimination to answer the query **P(G|e)** where E = {e, ~e) and the answer should be in the form of an expression. Eliminate the variables in the order the order A,B,C,D,F. **[12]**

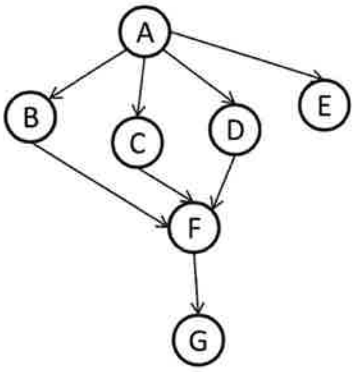


Figure (i)

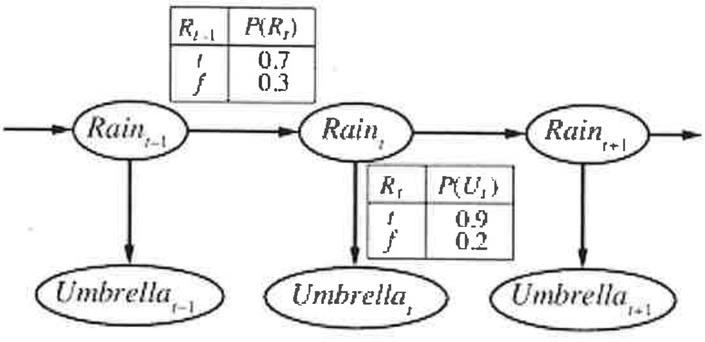


Figure (ii)

6. Consider the Bayesian Network in Figure (ii) and compute the smoothed estimate for probability of rain at time k = 1, given the umbrella observations to be true on day 1 and day 2. **[12]**

Q1 (i) Justification

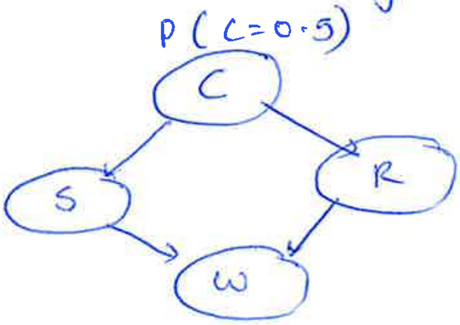
If the environment can change while an agent is deliberating (while the agent is deciding on the action), then the environment is considered as dynamic for the agent. Dynamic environment changes with the passage of time.

- Crossword Puzzle and similarly Tic-tac-toe game are static environment because when the agent is thinking to take action the environment (board configuration) does not change, on the hand for the other two tasks the environment is dynamic.

— No marks for this —

Q2 Likelihood weighting assigns weights to the samples depending on the likelihood of the evidence. The weight is the product of the conditional probabilities for the evidence variables given their parents. As the number of evidence variables increases most of the samples will have very low weights. However, there might still be a tiny fraction of samples that assign non-zero likelihood to the evidence. The weighted estimate will be dominated by this small fraction of samples.

Eg: consider the following network



Cloudy, sprinkler, Rain wet grass

C	P(S|C)
t	0.10
f	0.50

C	P(R|C)
t	0.8
f	.20

S R	P(W| S,r)
t t	.99
t f	.90
f t	.90
f f	.00

Consider the evidence variables +S, -W : P(- \|+s, -w)

The samples like the following will have very low weight as they are rare.

$+C \quad +S \quad +R \quad -W \; : \; 0.10 \times 0.01 = 0.001$

$+C \quad +S \quad +R \quad -W \; : \; 0.10 \times 0.01 = 0.001$

$\vdots$

$$\begin{cases} +C \quad +S \quad -R \quad -W \; : \; 0.10 \times 0.1 = 0.01 \\ -C \quad +S \quad -R \quad -W \; : \; 0.50 \times 0.1 = 0.05 \end{cases}$$

However, these samples influences the weighted estimates.

* justification with example = 2.5 + 1.5 = 4 marks

Q3 MH algorithm involves two distributions

proposal distribution : $q(\cdot)$

Target distribution : $P(x)$

* The assumption is that $P(x) = \frac{\tilde{P}(x)}{Z}$ where Z is the normalization constant, then $\tilde{P}(x)$ can be evaluated for any given value of x, although the value of Z may be unknown

* we can show that $P(x)$ is the stationary distribution of the Markov chain defined by the M-H algorithm by showing that the detailed balance is satisfied.

Let the current state be x and candidate state be x'

The transition probability can be given as

$$P(x \to x') = q(x' | x)\, \alpha(x' | x)$$

this is the acceptance prob!

in MH algorithm the x' is generated from proposal distribution

Proving detailed balance means showing that the flow from x to x', $\tilde{P}(x)\, P(x \to x')$ matches the flow from x' to x, $\tilde{P}(x')\, P(x' \to x)$.

$$\tilde{P}(x)\, P(x \to x') = \tilde{P}(x')\, P(x' \to x)$$

$$\tilde{p}(x)\ P(x \to x')$$

$$= \tilde{p}(x)\ q(x' | x)\ \alpha(x' | x)$$

$$= \tilde{p}(x)\ q(x' | x)\ \min\left(1, \frac{\tilde{p}(x')\ q(x | x')}{\tilde{p}(x)\ q(x' | x)}\right)$$

$$= \min\left(\tilde{p}(x)\ q(x' | x),\ \tilde{p}(x')\ q(x | x')\right)$$ multiplying in

$$= \tilde{p}(x')\ q(x | x')\ \min\left(\frac{\tilde{p}(x)\ q(x' | x)}{\tilde{p}(x')\ q(x | x')}, 1\right)$$ dividing out

$$= \tilde{p}(x')\ q(x | x')\ \alpha(x | x')$$

$$= \tilde{p}(x')\ P(x' \to x)$$

Hence proved.

$\boxed{Q4}$ (i) $P(Y=1 \mid W=1, X=0)$

2 marks $= 1 - 0.8 = 0.2$ (from the table)

(ii) The ordering is X, Z, W, Y.

In rejection sampling if the evidence appears later in the ordering then we may end up in rejecting many samples. In the given ordering X is sampled before Z as Z is the effect and X is the cause. Similarly, before sampling Y we need to sample both its parent X and W.

(iii) $W=0, X=0, Y=0, Z=1$ — weight $: 0.4$

(iv) $P(W=0 \mid X=0, Y=0, Z=1)$ Z=1 will be fixed as it is evidence

$$= \frac{P(W=0, X=0, Y=0)}{\sum_{w} P(W=w, X=0, Y=0)}$$

$$= \frac{0.5 \times 0.5 \times 0.6}{0.5 \times 0.5 \times 0.6 + 0.5 \times 0.5 \times 0.8} = 3/7 = 0.429$$

(i) $P(W=1 \mid X=0, Y=1)$

$$= \frac{\sum_{z} P(w, x, Y, z) \quad \text{—— ①}}{\sum_{w} \sum_{z} P(w, x, Y, Z) \quad \text{—— ②}}$$

① $= P(w)\,P(x)\,P(Y \mid w x)\,P(z \mid x) +$

$P(w)\,P(x)\,P(Y \mid w x)\,P(\sim z \mid x)$

$= 0.5 * 0.5 * 0.2 = 0.05$

②

$$= \sum_{w} \sum_{z} P(w)\,P(x)\,P(Y \mid w x)\,P(z \mid x)$$

$$= P(x) \sum_{w} P(w)\,P(Y \mid w x) \sum_{z} P(z \mid x)$$

$$= P(x)\left[P(w)\,P(Y \mid \overset{1\ 0}{w x}) + P(\sim w)\,P(Y \mid \overset{0\ 0}{\sim w x})\right]$$

$$= 0.5\left[0.5 * 0.2 + 0.5 * 0.4\right]$$

$$= \cancel{0.2}\,[0.15]$$

$$P(W=1 \mid X=0, Y=1) = \frac{0.05}{\cancel{0.2}\ 0.15} = \cancel{0.25}\ \cancel{0.25}\ \frac{1}{3}$$

Q5 $P(G|e) = \alpha P(G,e)$

$$= \alpha \sum_F \sum_D \sum_C \sum_B \sum_A P(A,B,C,D,e,F,G)$$

(2)

$$= \alpha \sum_F \sum_D \sum_C \sum_B \sum_A P(A)\,P(B|A)\,P(C|A)\,P(D|A)\,P(e|A)\,P(F|B,C,D)\,P(G|F)$$

(2) marks

$$= \alpha \sum_F \underbrace{P(G|F)}_{f_1(G,F)} \sum_D \sum_C \sum_B \underbrace{P(F|B,C,D)}_{f_2(F,B,C,D)} \sum_A \underbrace{P(A)}_{f_3(A)}\,\underbrace{P(B|A)}_{f_4(B,A)}\,\underbrace{P(C|A)}_{f_5(C,A)}\,\overbrace{P(D|A)}^{f_6(D,A)}\,\underbrace{P(e|A)}_{f_7(A)}$$

$$= \alpha \sum_F f_1(G,F) \sum_D \sum_C \sum_B f_2(F,B,C,D) \sum_A f_3(A)\,f_4(B,A)\,f_5(C,A)\,f_6(D,A)\,f_7(A)$$

eliminating A

$$f_8(B,C,D) = \sum_A f_3(A)\,f_4(B,A)\,f_5(C,A)\,f_6(D,A)\,f_7(A)$$

(2) marks

$$= f_3(a)\,f_4(B,a)\,f_5(C,a)\,f_6(D,a)\,f_7(a) + f_3(\sim a)\,f_4(B,\sim a)\,f_5(C,\sim a)\,f_6(D,\sim a)\,f_7(\sim a)$$

$$\alpha \sum_F f_1(G,F) \sum_D \sum_C \sum_B f_2(F,B,C,D)\,f_8(B,C,D)$$

eliminating B

$$f_9(F,D,C) = \sum_B f_2(F,B,C,D)\,f_8(B,C,D)$$

(2) marks

$$= f_2(F,b,C,D)\,f_8(b,C,D) + f_2(F,\sim b,C,D)\,f_8(\sim b,C,D)$$

eliminating C

$$f_{10}(F, D) = \sum_{C} f_9(F, D, c)$$

2

$$= f_9(F, D, c) + f_9(F, D, \sim c)$$

4 Marks

eliminating D

$$f_{11}(F) = \sum_{D} f_{10}(F, D)$$

2

$$= f_{10}(F, d) + f_{10}(f, \sim d)$$

eliminating F

$$f_{12}(G) = \sum_{F} f_1(G, F)\, f_{11}(F)$$

2 Marks

$$= f_1(G, f)\, f_{11}(f) + f_1(G, \sim f)\, f_{11}(\sim f)$$

finally we get,

$$P(G \mid e) = \alpha\, f_{12}(G)$$

Total 12 Marks