10-601: Homework 7

Due: 17 November 2014 11:59pm (Autolab)

TAs: Kuo Liu, Harry Gifford

Name: ____________________________

Andrew ID: ____________________________

Please answer to the point, and do not spend time/space giving irrelevant details. Please state any additional assumptions you make while answering the questions. For Questions 1 to 5, 6(b) and 6(c), you need to submit your answers in a single PDF file on autolab, either a scanned handwritten version or a LaTeXpdf file. Please make sure you write legibly for grading. For Question 6(a), submit your m-files on autolab.

You can work in groups. However, no written notes can be shared, or taken during group discussions. You may ask clarifying questions on Piazza. However, under no circumstances should you reveal any part of the answer publicly on Piazza or any other public website. The intention of this policy is to facilitate learning, not circumvent it. Any incidents of plagiarism will be handled in accordance with CMU's Policy on Academic Integrity.

⋆: Code of Conduct Declaration

- Did you receive any help whatsoever from anyone in solving this assignment? Yes / No.
- If you answered *yes*, give full details: ____________________________ (e.g. *Jane explained to me what is asked in Question 3.4*)
- Did you give any help whatsoever to anyone in solving this assignment? Yes / No.
- If you answered *yes*, give full details: ____________________________ (e.g. *I pointed Joe to section 2.3 to help him with Question 2*).

1: Warmup (TA:- Either)

Which of the following independence statements are true in the following directed graphical model? Explain your answer. $A \perp B \mid C$ can be read as "A is independent of B given C".

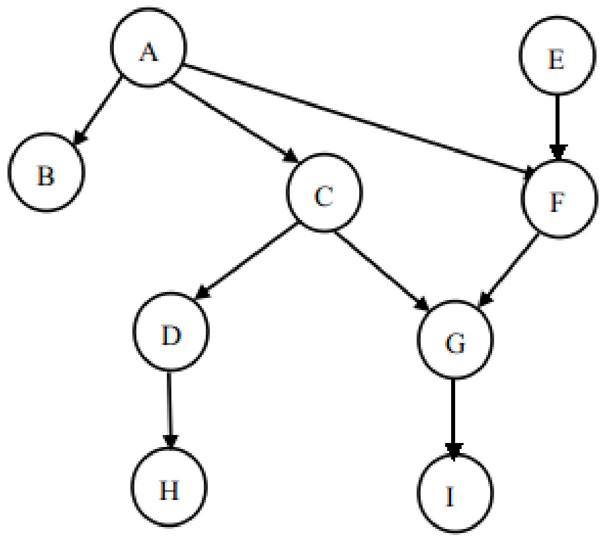


Figure 1: A simple DAG

(a) $A \perp E$

[3 points]

(b) $A \perp E \mid G$

[3 points]

(c) $C \perp F \mid \{A, G\}$

[3 points]

(d) $B \perp F \mid \{C, E\}$

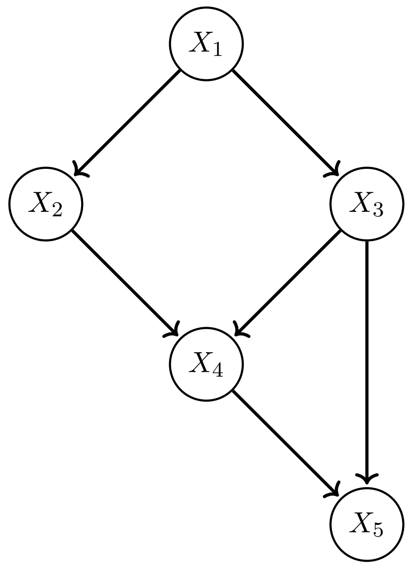
[3 points]

(e) $A \perp D \mid C$

[3 points]

2: Bayesian Networks (TA:- Harry Gifford)

(a) Prove that the any Bayesian network represents a valid probability distribution. Your proof should be general enough to apply to any graphical model, but to avoid clunky notation you may just prove it for the following graphical model. Specifically, you should show $\forall i, \forall X_i, \Pr(X_1, ..., X_n) \geq 0$ and $\sum_{X_1} \sum_{X_2} ... \sum_{X_n} \Pr(X_1, ..., X_n) = 1$.

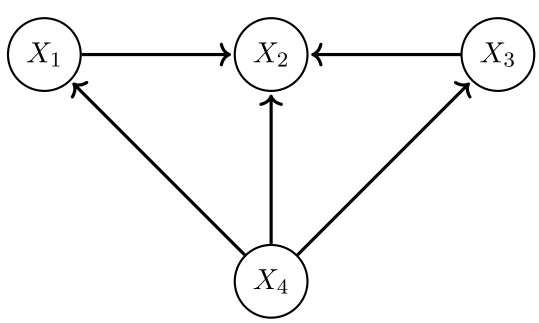


[*10 points*]

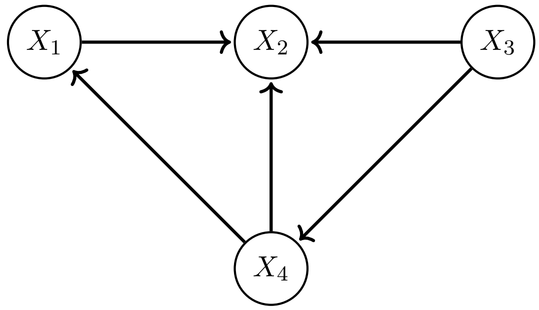
(b) Now we will explore the idea of equivalence for Bayesian networks. First some definitions. We say that two graphical models G_1 and G_2 are *Markov equivalent* if every independence statement in G_1 is also expressed in G_2 and likewise for G_2 into G_1. We define a *V-configuration*, $\langle i, j, k \rangle$ as a subgraph with three vertices and two edges connecting i, j and j, k. We say a V-configuration $\langle i, j, k \rangle$ is *shielded* if i is connected to k or k is connected to i. Finally, we say that two graphs have the same *skeleton* if the graphs obtained from removing the directions from the edges are the same.

For each pair of graphs below (i.e. (G_1, G_2), (G_1, G_3), (G_2, G_3)) state whether they are Markov equivalent or not. Explain your answer.

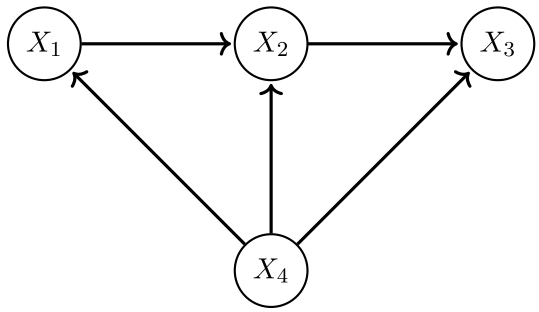
G_1



G_2



G_3



[5 points]

3: HMM I (TA:- Kuo Liu)

You have already learned forward method and backward method to compute the probability for a given observed sequence: $P(O_1 ... O_T)$
In this problem, we want to give you a different perspective of view to do this job and will use this new way to compute some probabilities for the following example HMM.

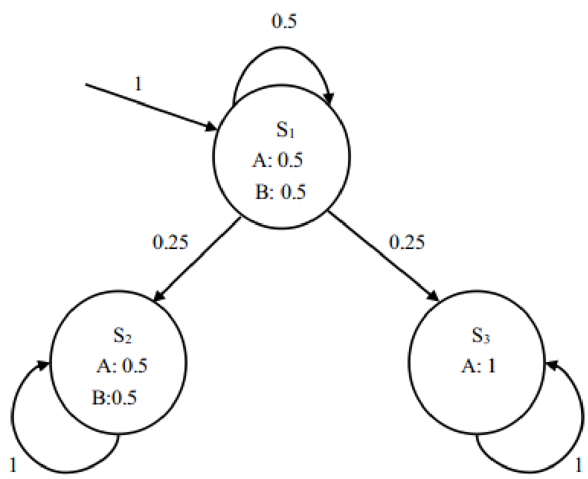


Figure 2: Figure of Q3, Initial and transition probabilities are listed next to the corresponding edges. Emission probabilities and the states' names are listed inside each node. For example, for state S_3 the emission probabilty is: 1.0 for A

(a) Let $v_i^t = p(O_1 ... O_T | q_t = s_i)$. Write a formula for $P(O_1 ... O_T)$ using **only** v_i^t and $p_t(i)$. [we define $p_t(i)$=$p(q_t = s_i)$]

[3 points]

(b) Compute $p(O_1 = B, ..., O_{200} = B)$ (the probability of observing 200 B's in the sequence)

[6 points]

(c) Compute $p(O_1 = A, ..., O_{200} = A)$ (the probability of observing 200 A's in the sequence)

[6 points]

4: HMM II (TA:- Kuo Liu)

Consider the HMM defined by the transition and emission probabilities in the table below. This HMM has six states (plus a start and end states) and an alphabet with four symbols (A,C,G and T). Thus, the probability of transitioning from state S_1 to state S_2 is 1, and the probability of emitting A while in state $S1$ is 0.5.

	0	S_1	S_2	S_3	S_4	S_5	S_6	f	A	C	G	T
0	0	1	0	0	0	0	0	0				
S_1	0	0	1	0	0	0	0	0	0.5	0.3	0	0.2
S_2	0	0	0	0.3	0	0.7	0	0	0.1	0.1	0.2	0.6
S_3	0	0	0	0	1	0	0	0	0.2	0	0.1	0.7
S_4	0	0	0	0	0	0	0	1	0.1	0.3	0.4	0.2
S_5	0	0	0	0	0	0	1	0	0.1	0.3	0.3	0.3
S_6	0	0	0	0	0	0	0	1	0.2	0.3	0	0.5

State whether the following are true or false and explain your answer.

(a) $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A, q_1 = S_1, q_2 = S_2) =$
$P(O_1 = A, O_2 = C, O_3 = T, O_4 = A | q_1 = S_1, q_2 = S_2)$

[*4 points*]

(b) $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A, q_3 = S_3, q_4 = S_4) >$
$P(O_1 = A, O_2 = C, O_3 = T, O_4 = A | q_3 = S_3, q_4 = S_4)$

[*4 points*]

(c) $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A, q_3 = S_3, q_4 = S_4) <$
$P(O_1 = A, O_2 = C, O_3 = T, O_4 = A | q_3 = S_5, q_4 = S_6)$

[*4 points*]

(d) $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A) > P(O_1 = A, O_2 = C, O_3 = T, O_4 = A, q_3 = S_3, q_4 = S_4)$

[*4 points*]

(e) $P(O_1 = A, O_2 = C, O_3 = T, O_4 = A) > P(O_1 = A, O_2 = C, O_3 = T, O_4 = A | q_3 = S_3, q_4 = S_4)$

[*4 points*]

5: HMM II (TA:- Kuo Liu)

Consider the following two state HMM, answer the following questions. You can either get the answer by hand calculation or write a program to get the final answer.

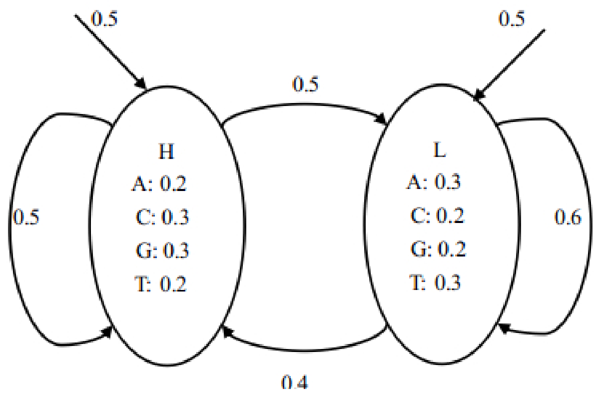


Figure 3: Figure of Q5

(a) What is the probability for you to get an output sequence like **GGCA**

[7 points]

(b) What is the most likely hidden status sequence for the output sequence **GGCACTGAA**

[8 points]