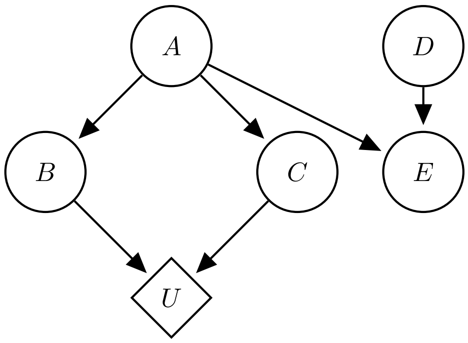
CS188: Exam Practice Session 9

Q1. VPI

(a) Consider a decision network with the following structure, where node U is the utility:



(i) For each of the following, choose the most specific option that is guaranteed to be true:

○ $VPI(B) = 0$ ○ $VPI(B) \geq 0$ ○ $VPI(B) > 0$

○ $VPI(D) = 0$ ○ $VPI(D) \geq 0$ ○ $VPI(D) > 0$

○ $VPI(E) = 0$ ○ $VPI(E) \geq 0$ ○ $VPI(E) > 0$

○ $VPI(A|E) = 0$ ○ $VPI(A|E) \geq 0$ ○ $VPI(A|E) > 0$

○ $VPI(E|A) = 0$ ○ $VPI(E|A) \geq 0$ ○ $VPI(E|A) > 0$

○ $VPI(A|B,C) = 0$ ○ $VPI(A|B,C) \geq 0$ ○ $VPI(A|B,C) > 0$

(ii) For each of the following, fill in the blank with the most specific of $>, \geq, <, \leq, =$ to guarantee that the comparison is true, or write ? if there is no possible guarantee.

$VPI(B)$ ________ $VPI(A)$

$VPI(B,C)$ ________ $VPI(A)$

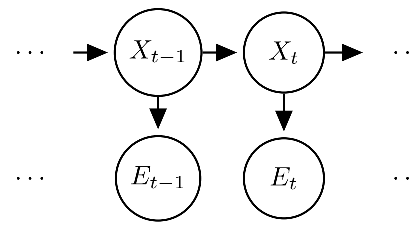
$VPI(B,C)$ ________ $VPI(B) + VPI(C)$

$VPI(B|C)$ ________ $VPI(A|C)$

Q2. HMMs - Forward Algorithm

Below is the forward algorithm update equation for Hidden Markov Models. As seen in lecture, we used $e_{1:t}$ to denote all the evidence variables $e_1, e_2, \ldots, e_t$. Similarly, $e_{1:t-1}$ denotes $e_1, e_2, \ldots, e_{t-1}$. For reference, the Bayes net corresponding to the usual Hidden Markov Model is shown on the right side of the equation below.

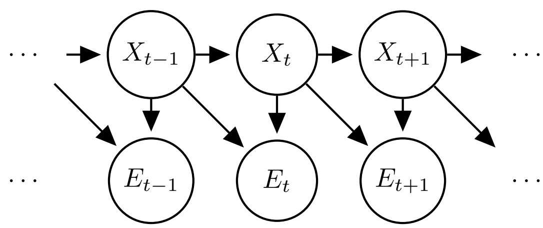
$$
\begin{aligned}
P(x_t|e_{1:t}) &\propto P(x_t, e_{1:t}) && (1)\\
&= P(e_t \mid x_t) \sum_{x_{t-1}} P(x_t \mid x_{t-1}) P(x_{t-1}, e_{1:t-1}) && (2)
\end{aligned}
$$



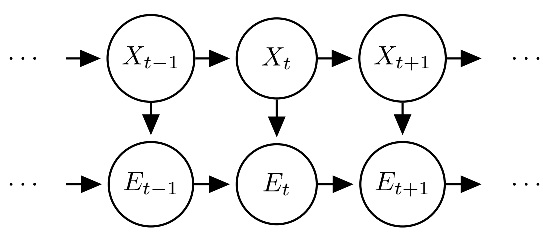
Hidden Markov Models can be extended in a number of ways to incorporate additional relations. Since the independence assumptions are different in these extended Hidden Markov Models, the forward algorithm updates will also be different.

Complete the forward algorithm updates for the extended HMMs specified by the following Bayes nets:

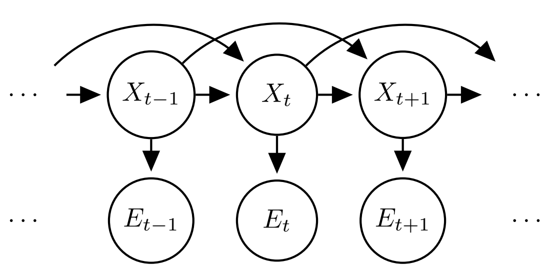
(a) $P(x_t|e_{1:t}) \propto \sum_{x_{t-1}} P(x_{t-1}, e_{1:t-1}) \cdot$ ________________________________



(b) $P(x_t|e_{1:t}) \propto \sum_{x_{t-1}} P(x_{t-1}, e_{1:t-1}) \cdot$ ________________________________



(c) $P(x_t, x_{t+1}|e_{1:t}) \propto \sum_{x_{t-1}} P(x_{t-1}, x_t, e_{1:t-1}) \cdot$ ________________________________

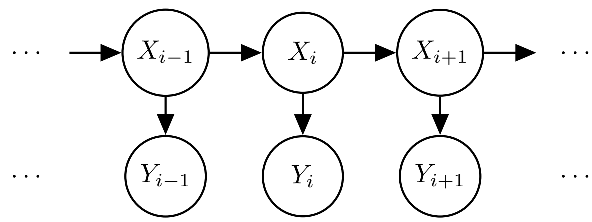


Q3. DNA Sequencing

Suppose you want to model the problem of DNA sequencing using the following set-up:

- $X_i, Y_i \in \{A, T, C, G\}$
- X_i : ith base of an individual
- Y_i : ith base output by DNA sequencer

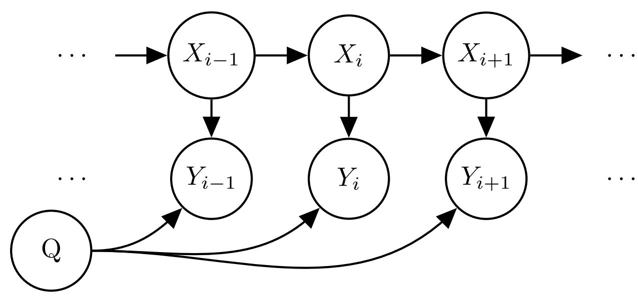
(a) First, you start by using a standard HMM model, shown below.



(i) Which of the following assumptions are made by the above HMM model

- ☐ $X_i \perp\!\!\!\perp X_j \quad \forall\, i \neq j$
- ☐ $Y_i \perp\!\!\!\perp Y_j \quad \forall\, i \neq j$
- ☐ $X_i \perp\!\!\!\perp Y_j \quad \forall\, i \neq j$
- ☐ $X_{i-1} \perp\!\!\!\perp X_{i+1} \mid X_i \quad \forall\, i$
- ☐ $X_i \perp\!\!\!\perp Y_{i+1} \mid X_{i+1} \quad \forall\, i$
- ☐ None of the provided options.

(b) Now you want to model the quality of your sequencer with a random variable Q, and decide to use the following modified HMM:



(i) Which of the following assumptions are made by the above modified HMM model?

- ☐ $X_i \perp\!\!\!\perp X_j \quad \forall\, i \neq j$
- ☐ $Y_i \perp\!\!\!\perp Y_j \quad \forall\, i \neq j$
- ☐ $X_i \perp\!\!\!\perp Y_j \quad \forall\, i \neq j$
- ☐ $X_{i-1} \perp\!\!\!\perp X_{i+1} \mid X_i \quad \forall\, i$
- ☐ $X_i \perp\!\!\!\perp Y_{i+1} \mid X_{i+1} \quad \forall\, i$
- ☐ $Q \perp\!\!\!\perp X_i \quad \forall\, i$
- ☐ $Q \perp\!\!\!\perp X_i \mid Y_i \quad \forall\, i$
- ☐ $Q \perp\!\!\!\perp X_i \mid Y_1, ...Y_N \quad \forall\, i$
- ☐ None of the provided options.

(ii) You observe the sequencer output $y_1, \ldots, y_N$ and want to estimate probability distribution of the particular sequence of length c starting at base k: $P(X_k \ldots X_{k+c-1} \mid y_1, \ldots y_N)$.
Select all elimination orderings which are maximally efficient with respect to the sum of the generated factors' sizes.

- ☐ $X_1, \ldots, X_{k-1}, X_{k+c}, \ldots, X_N, Q$
- ☐ $X_1, \ldots, X_{k-1}, X_N, \ldots, X_{k+c}, Q$
- ☐ $Q, X_1, \ldots, X_{k-1}, X_{k+c}, \ldots, X_N$
- ☐ $Q, X_1, \ldots, X_{k-1}, X_N, \ldots, X_{k+c}$
- ☐ $X_1, \ldots, X_{k-1}, Q, X_{k+c}, \ldots, X_N$
- ☐ $X_1, \ldots, X_{k-1}, Q, X_N, \ldots, X_{k+c}$
- ☐ None of the provided options: ____________

(iii) How many entries are in the final conditional probability table $P(X_k, \ldots, X_{k+c-1} \mid y_1, \ldots, y_N)$? The answer takes the form a^b – what are a and b?

$a =$ ________________________________ $b =$ ________________________________