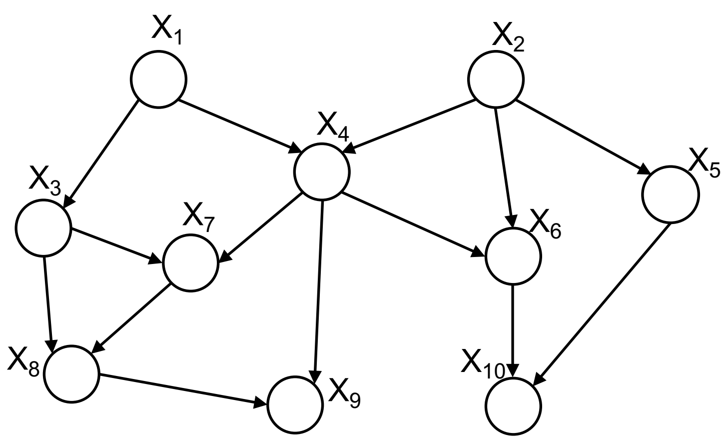
Written Assignment 4: Solution

Deadline: March 06, 2020

Instruction: You may discuss these problems with classmates, but please complete the write-ups individually. (This applies to BOTH undergraduates and graduate students.) Remember the collaboration guidelines set forth in class: you may meet to discuss problems with classmates, but you may not take any written notes (or electronic notes, or photos, etc.) away from the meeting. Your answers must be **typewritten**, except for figures or diagrams, which may be hand-drawn. Please submit your answers (pdf format only) on **Canvas**.

Q1. Bayes Nets: Independence (20 points)

Consider the following Bayesian network with 10 variables $\{X_1, X_2, \ldots, X_{10}\}$.



Which of the following statements are true:

1. $X_6 \perp\!\!\!\perp X_1 \mid X_2, X_4$
   Answer. True

2. $X_6 \perp\!\!\!\perp X_9 \mid X_4$
   Answer. False

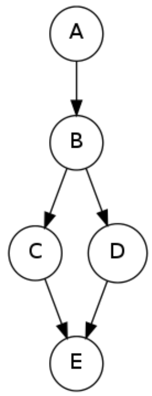
3. $X_3 \perp\!\!\!\perp X_9 \mid X_8$
   Answer. False

4. $X_1 \perp\!\!\!\perp X_2 \mid X_6$
   Answer. False

5. $X_4 \perp\!\!\!\perp X_8 \mid X_3, X_7$
   Answer. True

Q2. Bayes Nets: Inference (30 points)

Assume the following Bayes Net and corresponding CPTs.



A	$P(A)$
0	0.200
1	0.800

B	A	$P(B\|A)$
0	0	0.400
1	0	0.600
0	1	0.200
1	1	0.800

C	B	$P(C\|B)$
0	0	0.600
1	0	0.400
0	1	0.600
1	1	0.400

D	B	$P(D\|B)$
0	0	0.800
1	0	0.200
0	1	0.600
1	1	0.400

E	C	D	$P(E\|C,D)$
0	0	0	0.200
1	0	0	0.800
0	1	0	0.600
1	1	0	0.400
0	0	1	0.800
1	0	1	0.200
0	1	1	0.800
1	1	1	0.200

Compute the following conditional probabilities:

1. $P(B = 1 \mid E = 1)$
   Answer. Hidden variables are A, C, D. We will perform variable elimination using this order.

$$\text{Eliminate A: } f_1(B) = \sum_a P(a)P(B \mid a)$$

$$\text{Eliminate C: } f_2(B, D, E = 1) = \sum_c P(c \mid B)P(E = 1 \mid C, D)$$

$$\text{Eliminate D: } f_3(B, E = 1) = \sum_d P(d \mid B)f_2(B, d, E = 1)$$

$$\text{Join all remaining factors: } f_4(B, E = 1) = f_1(B) \times f_3(B, E = 1)$$

$$\text{Normalization: } P(B = 1 \mid E = 1) = \frac{f_4(B = 1, E = 1)}{f_4(B = 0, E = 1) + f_4(B = 1, E = 1)}$$

We obtain the following tables:

B	$f_1(B)$
0	0.24
1	0.76

B	D	E = 1	$f_2(B, D, E = 1)$
0	0	1	0.64
0	1	1	0.2
1	0	1	0.64
1	1	1	0.2

B	E = 1	$f_3(B, E = 1)$
0	1	0.552
1	1	0.464

B	E = 1	$f_4(B, E = 1)$
0	1	0.13248
1	1	0.35264

Therefore, $P(B = 1 \mid E = 1) = 0.727$

2. $P(A = 1 \mid C = 0, E = 0)$
Answer. Hidden variables are B, D. We will perform variable elimination using this order.

Eliminate B: $f_1(A, C = 0, D) = \sum_b P(b \mid A) P(C = 0 \mid b) P(D \mid b)$

Eliminate D: $f_2(A, C = 0, E = 0) = \sum_d f_1(A, C = 0, d) P(E = 0 \mid C = 0, d)$

Join all remaining factors: $f_3(A, C = 0, E = 0) = P(A) f_2(A, C = 0, E = 0)$

Normalization: $P(A = 1 \mid C = 0, E = 0) = \dfrac{f_3(A = 1, C = 0, E = 0)}{f_3(A = 0, C = 0, E = 0) + f_3(A = 1, C = 0, E = 0)}$

We obtain the following tables:

A	C = 0	D	$f_1(A, C = 0, D)$
0	0	0	0.408
0	0	1	0.192
1	0	0	0.384
1	0	1	0.216

A	C = 0	E= 0	$f_2(A, C = 0, E = 0)$
0	0	0	0.2352
1	0	0	0.2496

A	C = 0	E= 0	$f_3(A, C = 0, E = 0)$
0	0	0	0.04704
1	0	0	0.19968

Therefore, $P(A = 1 \mid C = 0, E = 0) = 0.809$

Q3. Bayes Nets: Sampling (10 points)

In this question, we will work with the same Bayes net and CPTs as **Q2**.

0.320	0.037	0.303	0.318	0.032	0.969	0.018	0.058	0.908	0.249

Q3.1. Rejection Sampling In this question, we will perform rejection sampling to estimate $P(C = 1 \mid B = 1, E = 1)$. Perform one round of rejection sampling, using the random samples given in the table below. Variables are sampled in the order A, B, C, D, E.

Note that the sampling attempt should stop as soon as you discover that the sample will be rejected. In that case mark the assignment of that variable and write "none" for the rest of the variables. When generating random samples, use as many values as needed from the table below, which we generated independently and uniformly at random from $[0, 1)$. Use numbers from left to right. To sample a binary variable W with probability $P(W = 0) = p$ and $P(W = 1) = 1 - p$ using a value a from the table, choose $W = 0$ if $a < p$ and $W = 1$ if $a \geq p$.

Choose the value (0 or 1) that each variable gets assigned to:

- A: 1 B: 0 C: none D: none E: none
- Which variable will get rejected? B.
  In rejection sampling, you reject any sample for which the variables' values do not match the values of the evidence variables in what you are trying to estimate. In this case, any sample where $B \neq 1$ or $e \neq 1$ is rejected. Only B and E can ever be rejected, in this case B was rejected because its sampled value was 0.

Q3.2. Likelihood Weighting In this question, we will perform likelihood weighting to estimate $P(C = 1 \mid B = 1, E = 1)$. Generate a sample and its weight, using the random samples given in the table below. Variables are sampled in the order A, B, C, D, E.

When generating random samples, use as many values as needed from the table below, which we generated independently and uniformly at random from $[0, 1)$. Use numbers from left to right. To sample a binary variable W with probability $P(W = 0) = p$ and $P(W = 1) = 1 - p$ using a value a from the table, choose $W = 0$ if $a < p$ and $W = 1$ if $a \geq p$.

0.249	0.052	0.299	0.773	0.715	0.550	0.703	0.105	0.236	0.153

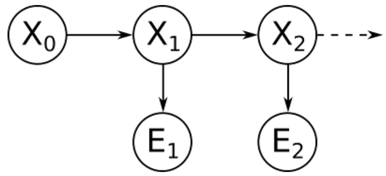
Select the assignments to the variables you sampled.

- A: 1 B: 1 C: 0 D: 0 E: 1
- What is the weight for the sample you obtained above? 0.64
  For likelihood weighting, the evidence variables, in this case B and E, are fixed and the sample is given weight equal to the product of the probabilities of the evidence variables taking on those values given the sampled values for their parents.

  In this case, the weight is equal to: $P(B = 1 \mid A = 1)P(E = 1 \mid C = 0, D = 0) = 0.64$.

Q4. Hidden Markov Models: Stationary Distribution (20 points)

Consider the HMM shown below:



The prior probability $P(X_0)$, dynamics model $P(X_{t+1} \mid X_t)$, and sensor model $P(E_t \mid X_t)$ are as follows. Provide the stationary distribution of X.

X_0	$P(X_0)$
0	0.2
1	0.8

X_{t+1}	X_t	$P(X_{t+1}\|X_t)$
0	0	0.3
1	0	0.7
0	1	0.05
1	1	0.95

E_t	X_t	$P(E_t\|X_t)$
a	0	0.3
b	0	0.15
c	0	0.55
a	1	0.1
b	1	0.45
c	1	0.45

Answer. We have the following linear equation system:

$$P_\infty(X = 0) = 0.3 \times P_\infty(X = 0) + 0.05 \times P_\infty(X = 1)$$
$$P_\infty(X = 1) = 0.7 \times P_\infty(X = 0) + 0.95 \times P_\infty(X = 1)$$
$$P_\infty(X = 0) + P_\infty(X = 1) = 1$$

The first and the second linear equations are both equivalent to $P_\infty(X = 1) = 14 \times P_\infty(X = 0)$. By combining it with the third equation, we obtain: $P_\infty(X = 1) \approx 0.933$ and $P_\infty(X = 0) \approx 0.067$.

Q5. Hidden Markov Models: The Forward Algorithm (20 points)

Consider the same HMM as **Q4** but with different probabilities as shown below:

X_0	$P(X_0)$
0	0.15
1	0.85

X_{t+1}	X_t	$P(X_{t+1}\|X_t)$
0	0	0.6
1	0	0.4
0	1	0.9
1	1	0.1

E_t	X_t	$P(E_t\|X_t)$
a	0	0.8
b	0	0.15
c	0	0.05
a	1	0.35
b	1	0.05
c	1	0.6

We have evidences $E_1 = E_2 = c$. What is the updated belief $B_2(X) = P(X_2 \mid E_1 = c, E_2 = c)$?

Answer. We apply the Forward algorithm:
First, the base case:

$$P(X_1 = 0) = P(X_1 = 0 \mid X_0 = 0)P(X_0) + P(X_1 = 0 \mid X_0 = 1)P(X_0 = 1)$$
$$= 0.6 \times 0.15 + 0.9 \times 0.85 = 0.855$$
$$P(X_1 = 1) = P(X_1 = 1 \mid X_0 = 0)P(X_0) + P(X_1 = 1 \mid X_0 = 1)P(X_0 = 1)$$
$$= 0.4 \times 0.15 + 0.1 \times 0.85 = 0.145$$
$$P(X_1 = 0 \mid E_1 = c) \propto P(X_1 = 0)P(E_1 = c \mid X_1 = 0) = 0.855 \times 0.05 = 0.04275$$
$$P(X_1 = 1 \mid E_1 = c) \propto P(X_1 = 1)P(E_1 = c \mid X_1 = 1) = 0.145 \times 0.6 = 0.087$$
$$\implies P(X_1 = 0 \mid E_1 = c) = \frac{0.04275}{0.04275 + 0.087} \approx 0.3295$$
$$\implies P(X_1 = 1 \mid E_1 = c) = \frac{0.087}{0.04275 + 0.087} \approx 0.6705$$

Next, passage of time update:

$$P(X_2 = 0 \mid E_1 = c)$$
$$= P(X_2 = 0 \mid X_1 = 0)P(X_1 = 0 \mid E_1 = c) + P(X_2 = 0 \mid X_1 = 1)P(X_1 = 1 \mid E_1 = c)$$
$$= 0.6 \times 0.3295 + 0.9 \times 0.6705 = 0.80115$$
$$P(X_2 = 1 \mid E_1 = c)$$
$$= P(X_2 = 1 \mid X_1 = 0)P(X_1 = 0 \mid E_1 = c) + P(X_2 = 1 \mid X_1 = 1)P(X_1 = 1 \mid E_1 = c)$$
$$= 0.4 \times 0.3295 + 0.1 \times 0.6705 = 0.19885$$

Finally, observation update:

$$P(X_2 = 0 \mid E_1 = c, E_2 = c) \propto P(E_2 = c \mid X_2 = 0)P(X_2 = 0 \mid E_1 = c) = 0.05 \times 0.80115 = 0.04$$
$$P(X_2 = 1 \mid E_1 = c, E_2 = c) \propto P(E_2 = c \mid X_2 = 1)P(X_2 = 1 \mid E_1 = c) = 0.6 \times 0.19885 = 0.11928$$
$$\implies P(X_2 = 0 \mid E_1 = c, E_2 = c) = \frac{0.04}{0.04 + 0.11928} \approx 0.2511$$
$$\implies P(X_2 = 1 \mid E_1 = c, E_2 = c) = 1 - 0.2511 \approx 0.7489$$