CS 188 Summer 2016 | Introduction to Artificial Intelligence

Practice Midterm II

To earn the extra credit, one of the following has to hold true. Please circle and sign.

A I spent 2 hours and 50 minutes or more on the practice midterm.

B I spent fewer than 2 hours and 50 minutes on the practice midterm, but I believe I have solved all the questions.

Signature: ____

The normal instructions for the midterm follow on the next page.

- You have approximately 2 Hours and 50 Minutes.
- The exam is closed book, closed notes except a one-page crib sheet.
- Please use non-programmable calculators only.
- Mark your answers ON THE EXAM ITSELF. If you are not sure of your answer you may wish to provide a *brief* explanation. All short answer sections can be successfully answered in a few sentences AT MOST.

First name	
Last name	
SID	
EdX username	

First and last name of student to your left	
First and last name of student to your right	

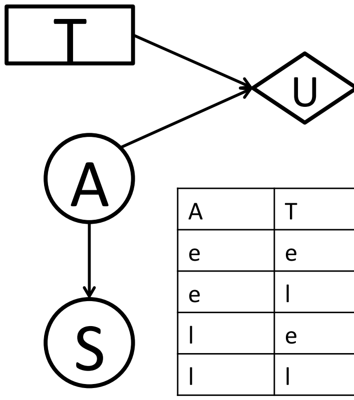
For staff use only:

Q1.	Probability and Decision Networks	/14
Q2.	Bayes' Nets: Elimination of a Single Variable	/11
Q3.	Probability	/10
Q4.	Bayes' Nets Representation and Probability	/10
Q5.	D-Separation	/16
Q6.	Independence in Hidden Markov Models	/12
Q7.	Bayes' Nets Sampling	/8
	Total	/81

For the questions that are not True/False, justify your answer concisely.

Q1. [14 pts] Probability and Decision Networks

A	P(A)
e	0.5
l	0.5



S	P(S)
e	
l	

A	S	P(S|A)
e	e	0.8
e	l	0.2
l	e	0.4
l	l	0.6

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

S	A	P(A|S)
e	e	
e	l	
l	e	
l	l	

Your parents are visiting you for graduation. You are in charge of picking them up at the airport. Their arrival time (A) might be early (e) or late (l). You decide on a time (T) to go to the airport, also either early (e) or late (l). Your sister (S) is a noisy source of information about their arrival time. The probability values and utilities are shown in the tables above.

Now we consider the case where you decide to ask your sister for input.

$EU(T = e|S = e) =$

$EU(T = l|S = e) =$

$MEU(\{S = e\}) =$

Optimal action with observation $\{S = e\}$ is $T =$

$EU(T = e|S = l) =$

$EU(T = l|S = l) =$

$MEU(\{S = l\}) =$

Optimal action with observation $S = l$ is $T =$

$VPI(S) =$

Q2. [11 pts] Bayes' Nets: Elimination of a Single Variable

Assume we are running variable elimination, and we currently have the following three factors:

A	B	$f_1(A,B)$
$+a$	$+b$	0.1
$+a$	$-b$	0.5
$-a$	$+b$	0.2
$-a$	$-b$	0.5

A	C	D	$f_2(A,C,D)$
$+a$	$+c$	$+d$	0.2
$+a$	$+c$	$-d$	0.1
$+a$	$-c$	$+d$	0.5
$+a$	$-c$	$-d$	0.1
$-a$	$+c$	$+d$	0.5
$-a$	$+c$	$-d$	0.2
$-a$	$-c$	$+d$	0.5
$-a$	$-c$	$-d$	0.2

B	D	$f_3(B,D)$
$+b$	$+d$	0.2
$+b$	$-d$	0.2
$-b$	$+d$	0.5
$-b$	$-d$	0.1

The next step in the variable elimination is to eliminate B.

(i) [3 pts] Which factors will participate in the elimination process of B?

(ii) [4 pts] Perform the join over the factors that participate in the elimination of B. Your answer should be a table similar to the tables above, it is your job to figure out which variables participate and what the numerical entries are.

(iii) [4 pts] Perform the summation over B for the factor you obtained from the join. Your answer should be a table similar to the tables above, it is your job to figure out which variables participate and what the numerical entries are.

Q3. [10 pts] Probability

(a) [2 pts] Select *all* of the expressions below that are equivalent to $P(A \mid B, C)$ given *no independence assumptions*.

- ○ $\sum_d P(A \mid B, C, D = d)$
- ○ $\sum_d P(A, D = d \mid B, C)$
- ○ $P(A \mid B)P(A \mid C)$
- ○ $P(A \mid C)$
- ○ $P(A \mid B)P(B \mid C)$
- ○ $P(A \mid C)P(C \mid B)$
- ○ $\frac{P(A,B,C)}{P(B,C)}$
- ○ $\frac{P(A)P(B|A)P(C|A,B)}{P(C)P(B|C)}$

(b) [2 pts] Select *all* of the expressions below that are equivalent to $P(A \mid B, C)$ given $A \perp\!\!\!\perp B$.

- ○ $\sum_d P(A \mid B, C, D = d)$
- ○ $\sum_d P(A, D = d \mid B, C)$
- ○ $P(A \mid B)P(A \mid C)$
- ○ $P(A \mid C)$
- ○ $P(A \mid B)P(B \mid C)$
- ○ $P(A \mid C)P(C \mid B)$
- ○ $\frac{P(A,B,C)}{P(B,C)}$
- ○ $\frac{P(A)P(B|A)P(C|A,B)}{P(C)P(B|C)}$

(c) [2 pts] Select *all* of the expressions below that are equivalent to $P(A \mid B, C)$ given $B \perp\!\!\!\perp C \mid A$.

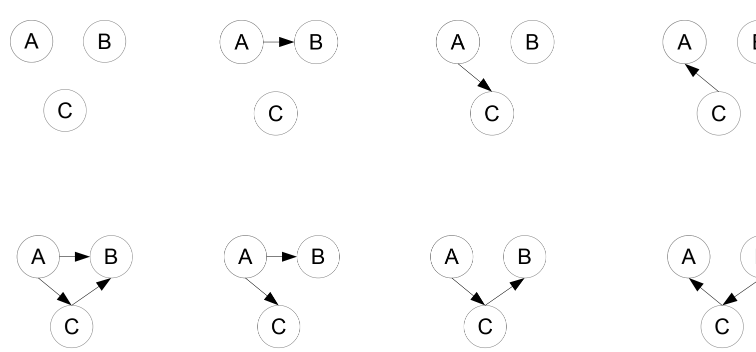
- ○ $\sum_d P(A \mid B, C, D = d)$
- ○ $\sum_d P(A, D = d \mid B, C)$
- ○ $P(A \mid B)P(A \mid C)$
- ○ $P(A \mid C)$
- ○ $P(A \mid B)P(B \mid C)$
- ○ $P(A \mid C)P(C \mid B)$
- ○ $\frac{P(A,B,C)}{P(B,C)}$
- ○ $\frac{P(A)P(B|A)P(C|A,B)}{P(C)P(B|C)}$

(d) [2 pts] Select *all* of the expressions below that hold for any distribution over four random variables A, B, C and D.

- ○ $P(A, B \mid C, D) = P(A \mid C, D)P(B \mid A, C, D)$
- ○ $P(A, B) = P(A, B \mid C, D)P(C, D)$
- ○ $P(A, B \mid C, D) = P(A, B)P(C, D)P(C, D \mid A, B)$
- ○ $P(A, B \mid C, D) = P(A, B)P(D)P(C, D \mid A, B)$

(e) [2 pts] Circle *all* of the Bayes' Nets below in which the following expression always holds:

$$P(A, B)P(C) = P(A)P(B, C)$$

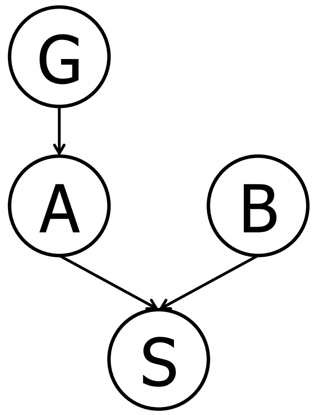


Q4. [10 pts] Bayes' Nets Representation and Probability

Suppose that a patient can have a symptom (S) that can be caused by two different diseases (A and B). It is known that the variation of gene G plays a big role in the manifestation of disease A. The Bayes' Net and corresponding conditional probability tables for this situation are shown below. For each part, you may leave your answer as an arithmetic expression.

$\mathbb{P}(G)$	
$+g$	0.1
$-g$	0.9

$\mathbb{P}(A \mid G)$		
$+g$	$+a$	1.0
$+g$	$-a$	0.0
$-g$	$+a$	0.1
$-g$	$-a$	0.9



$\mathbb{P}(B)$	
$+b$	0.4
$-b$	0.6

$\mathbb{P}(S \mid A, B)$			
$+a$	$+b$	$+s$	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

(a) [2 pts] Compute the following entry from the joint distribution:

$\mathbb{P}(+g, +a, +b, +s) =$

(b) [4 pts] What is the probability that a patient has disease A given that they have symptom S and disease B?

$\mathbb{P}(+a \mid +s, +b) =$

(c) [2 pts] What is the probability that a patient has the disease carrying gene variation G given that they have disease A?

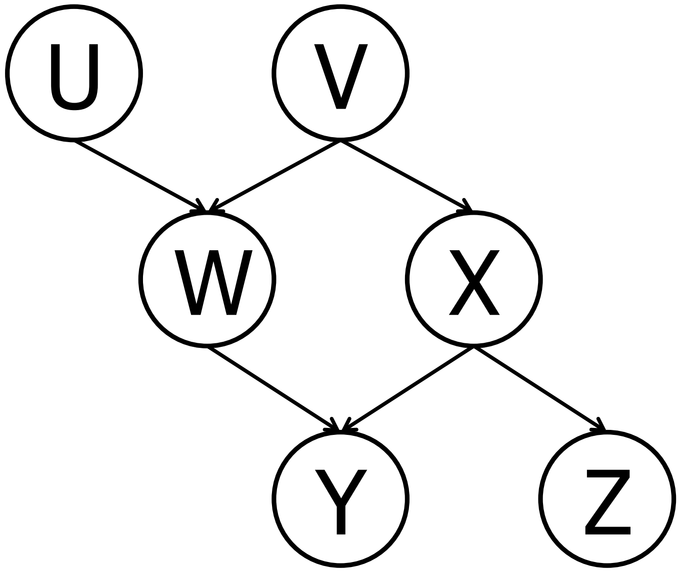
$\mathbb{P}(+g \mid +a) =$

(d) [2 pts] What is the probability that a patient has the disease carrying gene variation G given that they have disease B?

$\mathbb{P}(+g \mid +b) =$

Q5. [16 pts] D-Separation

(a) [16 pts] Based only on the structure of the (new) Bayes' Net given below, circle whether the following conditional independence assertions are guaranteed to be true, guaranteed to be false, or cannot be determined by the structure alone.

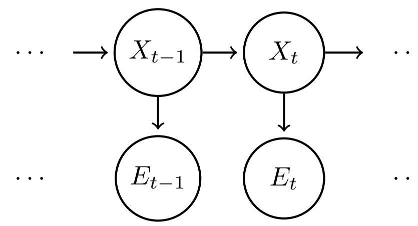


$U \perp\!\!\!\perp V$	Guaranteed true	Cannot be determined	Guaranteed false
$U \perp\!\!\!\perp V \mid W$	Guaranteed true	Cannot be determined	Guaranteed false
$U \perp\!\!\!\perp V \mid Y$	Guaranteed true	Cannot be determined	Guaranteed false
$U \perp\!\!\!\perp Z \mid W$	Guaranteed true	Cannot be determined	Guaranteed false
$U \perp\!\!\!\perp Z \mid V, Y$	Guaranteed true	Cannot be determined	Guaranteed false
$U \perp\!\!\!\perp Z \mid X, W$	Guaranteed true	Cannot be determined	Guaranteed false
$W \perp\!\!\!\perp X \mid Z$	Guaranteed true	Cannot be determined	Guaranteed false
$V \perp\!\!\!\perp Z \mid X$	Guaranteed true	Cannot be determined	Guaranteed false

Q6. [12 pts] Independence in Hidden Markov Models

Below is a full derivation of the forward algorithm updates 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 derivation below.

$$
\begin{aligned}
P(x_t|e_{1:t}) &\propto P(x_t, e_{1:t}) && (1)\\
&= \sum_{x_{t-1}} P(x_{t-1}, x_t, e_{1:t}) && (2)\\
&= \sum_{x_{t-1}} P(x_{t-1}, x_t, e_{1:t-1}, e_t) && (3)\\
&= \sum_{x_{t-1}} P(e_t \mid x_{t-1}, x_t, e_{1:t-1}) P(x_{t-1}, x_t, e_{1:t-1}) && (4)\\
&= \sum_{x_{t-1}} P(e_t \mid x_t) P(x_{t-1}, x_t, e_{1:t-1}) && (5)\\
&= \sum_{x_{t-1}} P(e_t \mid x_t) P(x_t \mid x_{t-1}, e_{1:t-1}) P(x_{t-1}, e_{1:t-1}) && (6)\\
&= \sum_{x_{t-1}} P(e_t \mid x_t) P(x_t \mid x_{t-1}) P(x_{t-1}, e_{1:t-1}) && (7)\\
&= P(e_t \mid x_t) \sum_{x_{t-1}} P(x_t \mid x_{t-1}) P(x_{t-1}, e_{1:t-1}) && (8)
\end{aligned}
$$



(a) [2 pts] The following assumption(s) are needed to justify going from step (4) to step (5):
(select all that apply)

- ◯ $E_t \perp\!\!\!\perp X_{t-1} \mid X_t$
- ◯ $E_t \perp\!\!\!\perp E_k \mid X_t$ for all $1 \le k \le t-1$
- ◯ $E_t \perp\!\!\!\perp E_k$ for all $1 \le k \le t-1$
- ◯ $E_t \perp\!\!\!\perp E_{t+1} \mid X_t$
- ◯ $E_t \perp\!\!\!\perp E_k \mid X_{t-1}$ for all $1 \le k \le t-1$
- ◯ $X_t \perp\!\!\!\perp E_{t+1} \mid X_{t+1}$
- ◯ $X_t \perp\!\!\!\perp E_k \mid X_{t-1}$ for all $1 \le k \le t-1$
- ◯ none

(b) [2 pts] The following assumption(s) are needed to justify going from step (5) to step (6):
(select all that apply)

- ◯ $E_t \perp\!\!\!\perp X_{t-1} \mid X_t$
- ◯ $E_t \perp\!\!\!\perp E_k \mid X_t$ for all $1 \le k \le t-1$
- ◯ $E_t \perp\!\!\!\perp E_k$ for all $1 \le k \le t-1$
- ◯ $E_t \perp\!\!\!\perp E_{t+1} \mid X_t$
- ◯ $E_t \perp\!\!\!\perp E_k \mid X_{t-1}$ for all $1 \le k \le t-1$
- ◯ $X_t \perp\!\!\!\perp E_{t+1} \mid X_{t+1}$
- ◯ $X_t \perp\!\!\!\perp E_k \mid X_{t-1}$ for all $1 \le k \le t-1$
- ◯ none

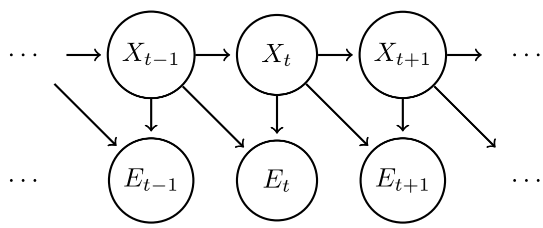
(c) [2 pts] The following assumption(s) are needed to justify going from step (6) to step (7):
(select all that apply)

- ◯ $E_t \perp\!\!\!\perp X_{t-1} \mid X_t$
- ◯ $E_t \perp\!\!\!\perp E_k \mid X_t$ for all $1 \le k \le t-1$
- ◯ $E_t \perp\!\!\!\perp E_k$ for all $1 \le k \le t-1$
- ◯ $E_t \perp\!\!\!\perp E_{t+1} \mid X_t$
- ◯ $E_t \perp\!\!\!\perp E_k \mid X_{t-1}$ for all $1 \le k \le t-1$
- ◯ $X_t \perp\!\!\!\perp E_{t+1} \mid X_{t+1}$
- ◯ $X_t \perp\!\!\!\perp E_k \mid X_{t-1}$ for all $1 \le k \le t-1$
- ◯ none

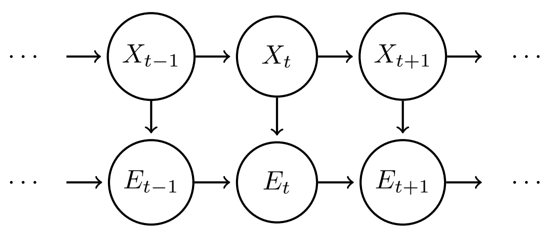
Hidden Markov Models can be extended in a number of ways to incorporate additional relations. Since the independence assumptions are different in these extended Hidden Markov Models, the forward algorithm updates will also be different.

Complete the forward algorithm updates for the extended Hidden Markov Models specified by the following Bayes nets:

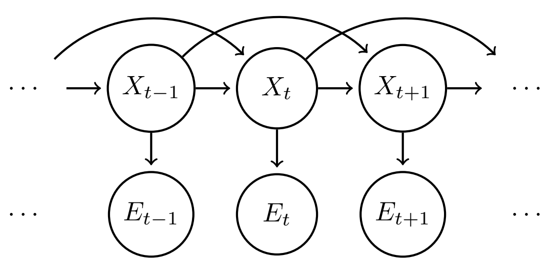
(d) [2 pts] $P(x_t|e_{1:t}) \propto \sum_{x_{t-1}} P(x_{t-1}, e_{1:t-1}) \cdot$ ____________________



(e) [2 pts] $P(x_t|e_{1:t}) \propto \sum_{x_{t-1}} P(x_{t-1}, e_{1:t-1}) \cdot$ ____________________

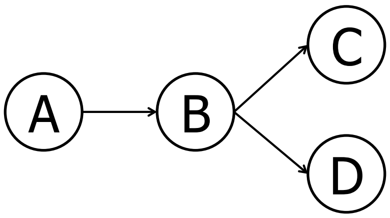


(f) [2 pts] $P(x_t, x_{t+1}|e_{1:t}) \propto \sum_{x_{t-1}} P(x_{t-1}, x_t, e_{1:t-1}) \cdot$ ____________________



Q7. [8 pts] Bayes' Nets Sampling

Assume the following Bayes net, and the corresponding distributions over the variables in the Bayes net:



A	$\mathbb{P}(A)$
$+a$	$1/5$
$-a$	$4/5$

A	B	$\mathbb{P}(B\|A)$
$+a$	$+b$	$1/5$
$+a$	$-b$	$4/5$
$-a$	$+b$	$1/2$
$-a$	$-b$	$1/2$

B	C	$\mathbb{P}(C\|B)$
$+b$	$+c$	$1/4$
$+b$	$-c$	$3/4$
$-b$	$+c$	$2/5$
$-b$	$-c$	$3/5$

B	D	$\mathbb{P}(D\|B)$
$+b$	$+d$	$1/2$
$+b$	$-d$	$1/2$
$-b$	$+d$	$4/5$
$-b$	$-d$	$1/5$

(a) [4 pts] Using the following samples (which were generated using likelihood weighting), estimate $\mathbb{P}(+b \mid -a, -c, -d)$ using likelihood weighting, or state why it cannot be computed.

$$\begin{array}{cccc} -a & -b & -c & -d \\ -a & +b & -c & -d \\ -a & -b & -c & -d \end{array}$$

(b) **(i)** [2 pts] Consider the query $P(A|-b,-c)$. After rejection sampling we end up with the following four samples: $(+a,-b,-c,+d), (+a,-b,-c,-d), (+a,-b,-c,-d), (-a,-b,-c,-d)$. What is the resulting estimate of $P(+a|-b,-c)$?

.

(ii) [2 pts] Consider again the query $P(A|-b,-c)$. After likelihood weighting sampling we end up with the following four samples: $(+a,-b,-c,-d), (+a,-b,-c,-d), (-a,-b,-c,-d), (-a,-b,-c,+d)$, and respective weights: $0.1, 0.1, 0.3, 0.3$. What is the resulting estimate of $P(+a|-b,-c)$?