CS 188 Summer 2020 | Introduction to Artificial Intelligence

Practice Midterm 2

- You have approximately 110 minutes.
- The exam is open book, open calculator, and open notes.
- In the interest of fairness, we want everyone to have access to the same information. To that end, we will not be answering questions about the content or making clarifications.
- For multiple choice questions,
  - ☐ means mark **all options** that apply
  - ◯ means mark a **single choice**

First name	
Last name	
SID	
Student to your right	
Student to your left	

For staff use only:

Q1.	Probability	/12
Q2.	Bayes Net Inference	/22
Q3.	HMMs: Help Your House Help You	/20
Q4.	Variable Elimination	/12
Q5.	Decision Networks and VPI	/21
Q6.	Bayes Nets Representation	/10
	Total	/97

THIS PAGE IS INTENTIONALLY LEFT BLANK

Q1. [12 pts] Probability

(a) A, B, C, and D are boolean random variables, and E is a random variable whose domain is $\{e_1, e_2, e_3, e_4, e_5\}$.

(i) [5 pts] How many entries are in the following probability tables and what is the sum of the values in each table? Write "?" if there is not enough information given.

Table	Size	Sum
$P(e \mid B)$		
$P(A, B \mid c)$		
$P(A, B \mid C, d, E)$		
$P(a, E \mid B, C)$		
$P(A, c, E)$		

(ii) [1 pt] What is the **minimum** number of parameters needed to fully specify the distribution $P(A, B|C, d, E)$

(iii) [1 pt] What is the **minimum** number of parameters needed to fully specify the distribution $P(a, E|B, C)$

(b) Given the same set of random variables as defined in part (a). Write each of the following expressions in its simplest form, i.e., a single term. Make no independence assumptions unless otherwise stated.

Write "*Not possible*" if it is not possible to simplify the expression without making further independence assumptions.

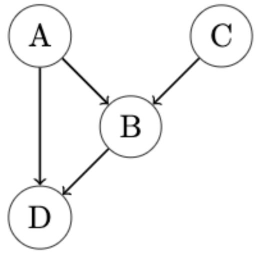
(i) [2 pts]

$$\sum_{a'} P(a' \mid B, E) P(c \mid a', B, E)$$

(ii) [3 pts]

$$\frac{\sum_{a'} P(B \mid a', C) P(a' \mid C) P(C)}{\sum_{d', e'} P(d' \mid e', C) P(e' \mid C) P(C)}$$

Q2. [22 pts] Bayes Net Inference



Consider the Bayes net graph depicted above.

(a) **(i)** [4 pts] Select all conditional independences that are enforced by this Bayes net graph.

☐ $A \perp\!\!\!\perp B$ ☐ $A \perp\!\!\!\perp C \mid B$ ☐ $D \perp\!\!\!\perp C \mid A, B$ ☐ $D \perp\!\!\!\perp C$

☐ $A \perp\!\!\!\perp C$ ☐ $A \perp\!\!\!\perp C \mid D$ ☐ $A \perp\!\!\!\perp C \mid B, D$ ☐ $D \perp\!\!\!\perp C \mid B$

(ii) [3 pts] Because of these conditional independences, there are some distributions that cannot be represented by this Bayes net. What is the minimum set of edges that would need to be added such that the resulting Bayes net could represent any distribution?

☐ $A \rightarrow C$ ☐ $C \rightarrow A$ ☐ $C \rightarrow D$ ☐ $D \rightarrow C$

☐ $D \rightarrow A$ ☐ $D \rightarrow B$ ☐ $B \rightarrow C$ ☐ $B \rightarrow A$

(b) [6 pts] For the rest of this Q2, we use the **original, unmodified** Bayes net depicted at the beginning of the problem statement. Here are some partially-filled conditional probability tables on A, B, C, and D. Note that these are not necessarily factors of the Bayes net. Fill in the six blank entries such that this distribution can be represented by the Bayes net.

A	C	$P(C \mid A)$
$+a$	$+c$	0.8
$+a$	$-c$	0.2
$-a$	$+c$	0.8
$-a$	$-c$	0.2

A	B	D	$P(D \mid A, B)$
$+a$	$+b$	$+d$	0.60
$+a$	$+b$	$-d$	0.40
$+a$	$-b$	$+d$	0.10
$+a$	$-b$	$-d$	0.90
$-a$	$+b$	$+d$	0.20
$-a$	$+b$	$-d$	0.80
$-a$	$-b$	$+d$	0.50
$-a$	$-b$	$-d$	0.50

A	B	C	$P(C \mid A, B)$
$+a$	$+b$	$+c$	0.50
$+a$	$+b$	$-c$	0.50
$+a$	$-b$	$+c$	0.20
$+a$	$-b$	$-c$	0.80
$-a$	$+b$	$+c$	0.90
$-a$	$+b$	$-c$	0.10
$-a$	$-b$	$+c$	0.40
$-a$	$-b$	$-c$	0.60

C	$P(C)$
$+c$	**(i)**
$-c$	**(ii)**

A	B	C	D	$P(D, C \mid A, B)$
$+a$	$+b$	$+c$	$+d$	**(iii)**
$+a$	$+b$	$-c$	$-d$	**(iv)**
$+a$	$-b$	$+c$	$+d$	**(v)**
$+a$	$-b$	$-c$	$-d$	**(vi)**
⋮	⋮	⋮	⋮	⋮

(i): ______ (ii): ______ (iii): ______

(iv): ______ (v): ______ (vi): ______

(c) Given the following conditional probability tables:

A	$P(A)$
$+a$	0.05
$-a$	0.95

C	$P(C)$
$+c$	0.3
$-c$	0.7

$P(B \mid A, C)$			
$+a$	$+c$	$+b$	0.65
$+a$	$+c$	$-b$	0.35
$+a$	$-c$	$+b$	0.15
$+a$	$-c$	$-b$	0.85
$-a$	$+c$	$+b$	0.25
$-a$	$+c$	$-b$	0.75
$-a$	$-c$	$+b$	0.55
$-a$	$-c$	$-b$	0.45

$P(D \mid A, B)$			
$+a$	$+b$	$+d$	0.60
$+a$	$+b$	$-d$	0.40
$+a$	$-b$	$+d$	0.10
$+a$	$-b$	$-d$	0.90
$-a$	$+b$	$+d$	0.20
$-a$	$+b$	$-d$	0.80
$-a$	$-b$	$+d$	0.50
$-a$	$-b$	$-d$	0.50

(i) [5 pts] Suppose that we want to use likelihood weighted sampling to approximate $P(A \mid +b, +c, +d)$. However, we accidentally forgot to fix the value of C and D, and instead we sampled them just like unconditioned variables!

For each of the samples below, write what the weight of the sample should be, in order to correctly approximate $P(A \mid +b, +c, +d)$. If the weight of the sample does not matter for calculating $P(A \mid +b, +c, +d)$, write 'reject' instead (since we would not use that sample).

$(+a, +b, -c, +d)$: [ ]

$(+a, +b, +c, +d)$: [ ]

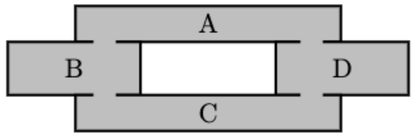
$(-a, +b, +c, +d)$: [ ]

(ii) [4 pts] Let's say we're trying to approximate $P(A \mid -b)$ using Gibbs sampling. Suppose the most recent sample is $(+a, -b, +c, +d)$ If we choose D to resample, what is the probability of resampling $+d$ and $-d$ respectively?

$+d$: [ ] $-d$: [ ]

Q3. [20 pts] HMMs: Help Your House Help You

Imagine you have a smart house that wants to track your location within itself so it can turn on the lights in the room you are in and make you food in your kitchen. Your house has 4 rooms (A, B, C, D) in the floorplan below (A is connected to B and D, B is connected to A and C, C is connected to B and D, and D is connected to A and C):



At the beginning of the day $(t = 0)$, your probabilities of being in each room are p_A, p_B, p_C, and p_D for rooms A, B, C, and D, respectively, and at each time t your position (following a Markovian process) is given by X_t. At each time, your probability of staying in the same room is q_0, your probability of moving clockwise to the next room is q_1, and your probability of moving counterclockwise to the next room is $q_{-1} = 1 - q_0 - q_1$.

(a) [3 pts] Initially, assume your house has no way of sensing where you are. What is the probability that you will be in room D at time $t = 1$?

○ $q_0 p_D$ ○ $q_0 p_D + q_1 p_A + q_{-1} p_C + 2q_1 p_B$ ○ $q_0 p_D + q_1 p_A + q_{-1} p_C$

○ $q_0 p_D + q_{-1} p_A + q_1 p_C$ ○ $q_1 p_A + q_1 p_C + q_0 p_D$ ○ None of these

Now assume your house contains a sensor M^A that detects motion ($+m$) or no motion ($-m$) in room A. However, the sensor is a bit noisy and can be tricked by movement in adjacent rooms, resulting in the conditional distributions for the sensor given in the table below. The prior distribution for the sensor's output is also given.

M^A	$P(M^A \mid X = A)$	$P(M^A \mid X = B)$	$P(M^A \mid X = C)$	$P(M^A \mid X = D)$
$+m^A$	$1 - 2\gamma$	γ	0.0	γ
$-m^A$	2γ	$1 - \gamma$	1.0	$1 - \gamma$

M^A	$P(M^A)$
$+m^A$	0.5
$-m^A$	0.5

(b) [3 pts] You decide to help your house to track your movements using a particle filter with three particles. At time $t = T$, the particles are at $X^0 = A, X^1 = B, X^2 = D$. What is the probability that the particles will be resampled as $X^0 = X^1 = X^2 = A$ after time elapse? Select **all terms in the product.**

☐ q_0 ☐ q_0^2 ☐ q_0^3 ☐ q_1 ☐ q_1^2 ☐ q_1^3 ☐ q_{-1} ☐ q_{-1}^2 ☐ q_{-1}^3 ○ None of these

(c) [3 pts] Assume that the particles are actually resampled after time elapse as $X^0 = D, X^1 = B, X^2 = C$, and the sensor observes $M^A = -m^A$. What are the particle weights given the observation?

Particle	Weight						
X^0 = D	○ γ	○ $1 - \gamma$	○ $1 - 2\gamma$	○ 0.0	○ 1.0	○ 2γ	○ None of these
X^1 = B	○ γ	○ $1 - \gamma$	○ $1 - 2\gamma$	○ 0.0	○ 1.0	○ 2γ	○ None of these
X^2 = C	○ γ	○ $1 - \gamma$	○ $1 - 2\gamma$	○ 0.0	○ 1.0	○ 2γ	○ None of these

Now, assume your house also contains sensors M^B and M^D in rooms B and D, respectively, with the conditional distributions of the sensors given below and the prior equivalent to that of sensor M^A.

M^B	$P(M^B \mid X = A)$	$P(M^B \mid X = B)$	$P(M^B \mid X = C)$	$P(M^B \mid X = D)$
$+m^B$	γ	$1 - 2\gamma$	γ	0.0
$-m^B$	$1 - \gamma$	2γ	$1 - \gamma$	1.0

M^D	$P(M^D \mid X = A)$	$P(M^D \mid X = B)$	$P(M^D \mid X = C)$	$P(M^D \mid X = D)$
$+m^D$	γ	0.0	γ	$1 - 2\gamma$
$-m^D$	$1 - \gamma$	1.0	$1 - \gamma$	2γ

(d) [6 pts] Again, assume that the particles are actually resampled after time elapse as $X^0 = D, X^1 = B, X^2 = C$. The sensor readings are now $M^A = -m^A, M^B = -m^B, M^D = +m^D$. What are the particle weights given the observations?

Particle	Weight
X^0 = D	○ $\gamma^2 - 2\gamma^3$ ○ $3 - 2\gamma$ ○ 0.0 ○ $\gamma - \gamma^2 + \gamma^3$ ○ $1 - 3\gamma + 2\gamma^2$ ○ $2 - \gamma$ ○ $1 - 2\gamma + \gamma^2$ ○ None of these
X^1 = B	○ $\gamma^2 - 2\gamma^3$ ○ $3 - 2\gamma$ ○ 0.0 ○ $\gamma - \gamma^2 + \gamma^3$ ○ $1 - 3\gamma + 2\gamma^2$ ○ $2 - \gamma$ ○ $1 - 2\gamma + \gamma^2$ ○ None of these
X^2 = C	○ $\gamma^2 - 2\gamma^3$ ○ $3 - 2\gamma$ ○ 0.0 ○ $\gamma - \gamma^2 + \gamma^3$ ○ $1 - 3\gamma + 2\gamma^2$ ○ $2 - \gamma$ ○ $1 - 2\gamma + \gamma^2$ ○ None of these

The sequence of observations from each sensor are expressed as the following: $m^A_{0:t}$ are all measurements $m^A_0, m^A_1, \ldots, m^A_t$ from sensor M^A, $m^B_{0:t}$ are all measurements $m^B_0, m^B_1, \ldots, m^B_t$ from sensor M^B, and $m^D_{0:t}$ are all measurements $m^D_0, m^D_1, \ldots, m^D_t$ from sensor M^D. Your house can get an accurate estimate of where you are at a given time t using the forward algorithm. The forward algorithm update step is shown here:

$$
\begin{aligned}
P(X_t \mid m^A_{0:t}, m^B_{0:t}, m^D_{0:t}) &\propto P(X_t, m^A_{0:t}, m^B_{0:t}, m^D_{0:t}) && (1)\\
&= \sum_{x_{t-1}} P(X_t, x_{t-1}, m^A_t, m^B_t, m^D_t, m^A_{0:t-1}, m^B_{0:t-1}, m^D_{0:t-1}) && (2)\\
&= \sum_{x_{t-1}} \_\_\_\_\_\_\_\_ P(X_t \mid x_{t-1}) P(x_{t-1}, m^A_{0:t-1}, m^B_{0:t-1}, m^D_{0:t-1}) && (3)
\end{aligned}
$$

(e) [5 pts] Which of the following expression(s) correctly complete the missing expression in line (3) above (regardless of whether they are available to the algorithm during execution)? Fill in **all** that apply.

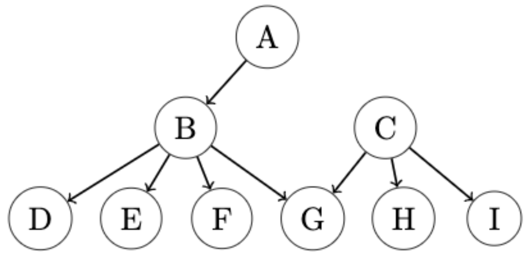
☐ $P(m^A_t, m^B_t, m^D_t \mid X_t, x_{t-1})$ ☐ $P(m^A_t, m^B_t, m^D_t \mid x_{t-1})$ ☐ $P(m^A_t \mid x_{t-1})P(m^B_t \mid x_{t-1})P(m^D_t \mid x_{t-1})$

☐ $P(m^A_t \mid m^A_{0:t-1})P(m^B_t \mid m^B_{0:t-1})P(m^D_t \mid m^D_{0:t-1})$ ☐ $P(m^A_t, m^B_t, m^D_t \mid X_t, x_{t-1}, m^A_{0:t-1}, m^B_{0:t-1}, m^D_{0:t-1})$

☐ $P(m^A_t \mid X_t)P(m^B_t \mid X_t)P(m^D_t \mid X_t)$ ☐ $P(m^A_t, m^B_t, m^D_t \mid X_t)$ ○ None of these

Q4. [12 pts] Variable Elimination

Consider the following Bayes Net:



(a) [4 pts] Given the following domain sizes for the variables:

Variable	**Domain Size**
A	2^2
B	2^3
C	2^8
D	2^5
E	2^6
F	2^7
G	2^8
H	2^9
I	2^{10}

What is the size of the biggest factor generated when we perform variable elimination with an alphabetical elimination order for the query $P(G = g \mid I = i)$ for some $g \in \text{dom}(G)$ and $i \in \text{dom}(I)$?

(b) [3 pts] Which is the variable whose elimination generates the biggest factor if we perform variable elimination in alphabetical order for the query $P(G = g \mid I = i)$ for some $g \in \text{dom}(G)$ and $i \in \text{dom}(I)$?

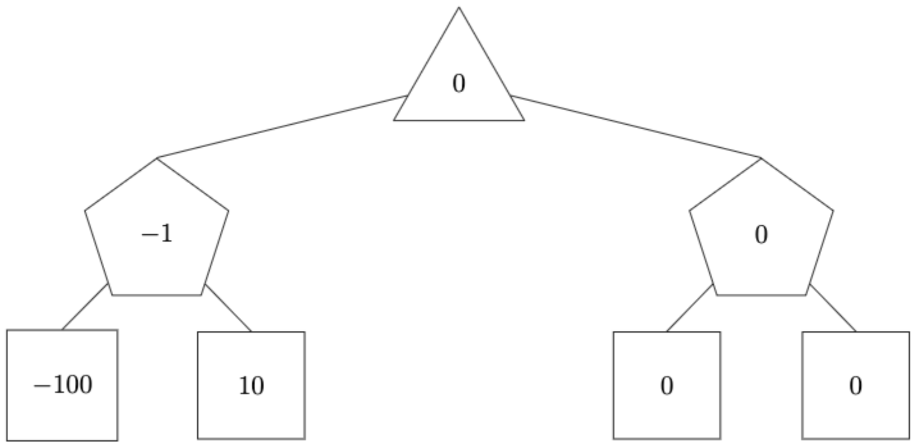
○ A ○ B ○ C ○ D ○ E ○ F ○ G ○ H ○ I
○ None of the above

(c) [5 pts] *Now, suppose the variables are all* ***boolean variables***, give an elimination ordering that generates the smallest largest factor for the query $P(A = a \mid I = i)$ for some $a \in \text{dom}(A)$ and $i \in \text{dom}(I)$.
Leave the extra boxes **blank**. For example, if your elimination ordering is X, Y, Z, you should only fill in the first 3 boxes.

Q5. [21 pts] Decision Networks and VPI

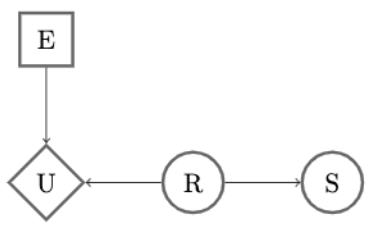
Valerie has just found a cookie on the ground. She is concerned that the cookie contains raisins, which she really dislikes but she still wants to eat the cookie. If she eats the cookie and it contains raisins she will receive a utility of −100 and if the cookie doesn't contain raisins she will receive a utility of 10. If she doesn't eat the cookie she will get 0 utility. The cookie contains raisins with probability 0.1.

(a) [1 pt] We represent this decision network as the following expectimax game tree.



Should Valerie eat the cookie? ○ Yes ○ No

(b) [4 pts] Valerie can now smell the cookie to judge whether it has raisins before she eats it. However, since she dislikes raisins she does not have much experience with them and cannot recognize their smell well. As a result she will incorrectly identify raisins when there are no raisins with probability 0.2 and will incorrectly identify no raisins when there are raisins with probability 0.3. This decision network can be represented by the diagram below where E is her choice to eat, U is her utility earned, R is whether the cookie contains raisins, and S is her attempt at smelling.



Valerie has just smelled the cookie and she thinks it doesn't have raisins. Write the probability, X, that the cookie has raisins given that she smelled no raisins as a simplest form fraction or decimal.

$X =$ ____________________

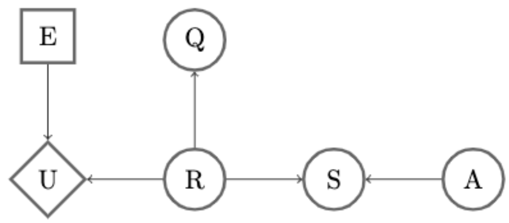
(c) [4 pts] What is her maximum expected utility, Y given that she smelled no raisins? You can answer in terms of X or as a simplest form fraction or decimal.

$Y =$ ____________________

(d) [4 pts] What is the Value of Perfect Information (VPI) of smelling the cookie? You can answer in terms of X and Y or as a simplest form fraction or decimal.

$VPI =$ [ ]

(e) [8 pts] Valerie is unsatisfied with the previous model and wants to incorporate more variables into her decision network. First, she realizes that the air quality (A) can affect her smelling accuracy. Second, she realizes that she can question (Q) the people around to see if they know where the cookie came from. These additions are reflected in the decision network below.



Choose one for each equation:

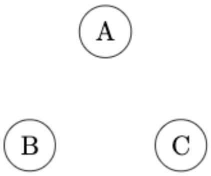
	Could Be True	Must Be True	Must Be False
$VPI(A, S) > VPI(A) + VPI(S)$	○	○	○
$VPI(A) = 0$	○	○	○
$VPI(Q, R) \leq VPI(Q) + VPI(R)$	○	○	○
$VPI(S, R) > VPI(R)$	○	○	○
$VPI(Q) \geq 0$	○	○	○
$VPI(Q, A) > VPI(Q)$	○	○	○
$VPI(S \mid A) < VPI(S)$	○	○	○
$VPI(A \mid S) > VPI(A)$	○	○	○

Q6. [10 pts] Bayes Nets Representation

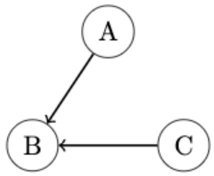
(a) [5 pts] Given the joint probability table on the right.

Clearly fill in **all** circles corresponding to Bayes Nets (BNs) that can correctly represent the distribution on the right. If no such Bayes Nets are given, clearly select *None of the above.*

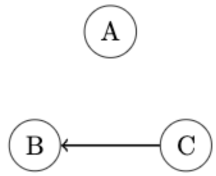
A	**B**	**C**	**P(A, B, C)**
0	0	0	.22
0	0	1	.08
0	1	0	.22
0	1	1	.08
1	0	0	.09
1	0	1	.11
1	1	0	.09
1	1	1	.11



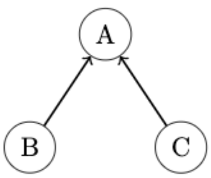
○ $\mathbf{G_1}$



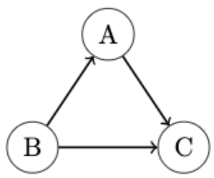
○ $\mathbf{G_2}$



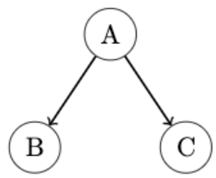
○ $\mathbf{G_3}$



○ $\mathbf{G_4}$



○ $\mathbf{G_5}$

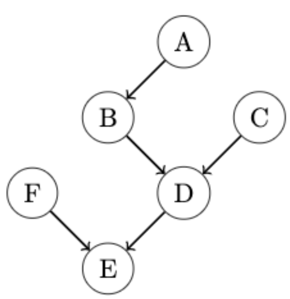


○ $\mathbf{G_6}$

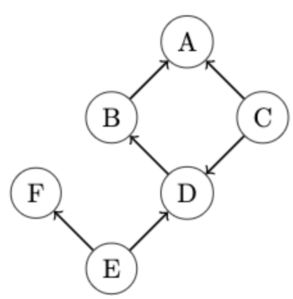
○ **None of the above.**

(b) [5 pts] For the pair of Bayes Net (BN) models below, indicate if the New BN model is guaranteed to be able to represent any joint distribution that the Old BN Model can represent. If the New BN model is guaranteed to be able to represent any joint distribution that the Old BN Model can represent, select "None." Otherwise, fill in the squares corresponding to the minimal number of edges that must be added such that the modified New BN can represent any joint distribution that the Old BN Model can represent. Select "Not Possible" if no combination of added edges can result in the modified New BN representing any joint distribution that the Old BN Model can represent.

Old BN Model



New BN Model



☐ $A \to D$ ☐ $A \to E$ ☐ $A \to F$ ☐ $B \to C$ ☐ $B \to E$ ☐ $B \to F$
☐ $C \to B$ ☐ $C \to E$ ☐ $C \to F$ ☐ $D \to A$ ☐ $D \to F$ ☐ $E \to A$
☐ $E \to B$ ☐ $E \to C$ ☐ $F \to A$ ☐ $F \to B$ ☐ $F \to C$ ☐ $F \to D$
○ None ○ Not Possible