Probability and Bayes Nets

(a) A, B, and C are random variables with binary domains. How many entries are in the following probability tables and what is the sum of the values in each table? Write a "?" in the box if there is not enough information given.

size 2^N — no. of var.

Table	Size	Sum
$P(A, B\|C)$	8	2
$P(A\| +b, +c)$	2	1
$P(+a\|B)$	2	?

Sum 1 for each distribution

(b) Circle true if the following probability equalities are valid and circle false if they are invalid (leave it blank if you don't wish to risk a guess). Each True/False question is worth 1 points. Leaving a question blank is worth 0 points. **Answering incorrectly is worth −1 points.**

No independence assumptions are made.

(i) [*true* or *false*] $P(A, B) = P(A|B)P(A)$ $P(B)$

(ii) [*true* or *false*] $P(A|B)P(C|B) = P(A, C|B)$ — $\frac{P(A,B,C)}{P(B)}$ no way

(iii) [*true* or *false*] $P(B, C) = \sum_{a \in A} P(B, C|A)$ sum on right side

(iv) [*true* or *false*] $P(A, B, C, D) = P(C)P(D|C)P(A|C, D)P(B|A, C, D)$ $P(B, A, D, C)$

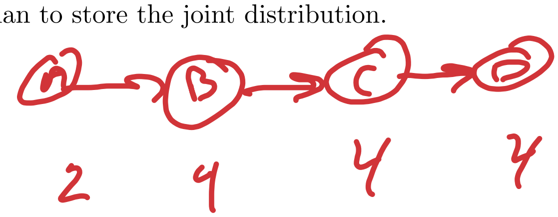
(c) Space Complexity of Bayes Nets

Consider a joint distribution over N variables. Let k be the domain size for all of these variables, and let d be the maximum indegree of any node in a Bayes net that encodes this distribution.

(i) What is the space complexity of storing the entire joint distribution? Give an answer of the form $O(\cdot)$.

$O(k^d)$

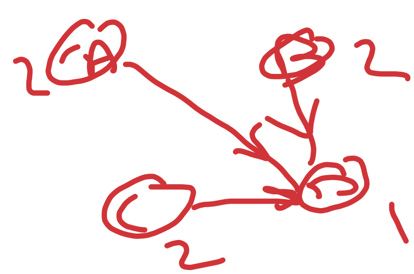
(ii) Draw an example of a Bayes net over four binary variables such that it takes less space to store the Bayes net than to store the joint distribution.



$\to$ 14 BN

$2^4 = 16$ joint Dis

(iii) Draw an example of a Bayes net over four binary variables such that it takes more space to store the Bayes net than to store the joint distribution.



22 BN

Probability

(a) For the following questions, you will be given a set of probability tables and a set of conditional independence assumptions. Given these tables and independence assumptions, write an expression for the requested probability tables. Keep in mind that your expressions cannot contain any probabilities other than the given probability tables. If it is not possible, mark "Not possible."

(i) Using probability tables $\mathbf{P(A), P(A \mid C), P(B \mid C), P(C \mid A, B)}$ and no conditional independence assumptions, write an expression to calculate the table $\mathbf{P(A, B \mid C)}$.

$\frac{P(A,B,C)}{P(C)}$ ← no way

$\mathbf{P(A, B \mid C)} =$ ____________________ - [x] Not possible.

(ii) Using probability tables $\mathbf{P(A), P(A \mid C), P(B \mid A), P(C \mid A, B)}$ and no conditional independence assumptions, write an expression to calculate the table $\mathbf{P(B \mid A, C)}$.

$\frac{P(A,B,C)}{P(A,C)}$

$\mathbf{P(B \mid A, C)} = \frac{P(C|A,B)\, P(B|A)\, P(A)}{\sum_B}$ - [ ] Not possible.

(iii) Using probability tables $\mathbf{P(A \mid B), P(B), P(B \mid A, C), P(C \mid A)}$ and conditional independence assumption $\mathbf{A \perp\!\!\!\perp B}$, write an expression to calculate the table $\mathbf{P(C)}$.

$P(A)$

$\mathbf{P(C)} = \sum_A P(A|B)\, P(C|A)$ - [ ] Not possible.

(iv) Using probability tables $\mathbf{P(A \mid B, C), P(B), P(B \mid A, C), P(C \mid B, A)}$ and conditional independence assumption $\mathbf{A \perp\!\!\!\perp B \mid C}$, write an expression for $\mathbf{P(A, B, C)}$.

$\mathbf{P(A, B, C)} =$ ____________________ - [x] Not possible.

(b) For each of the following equations, select the *minimal set* of conditional independence assumptions necessary for the equation to be true.

(i) $\mathbf{P(A, C) = \underline{P(A \mid B)}\ P(C)}$

$P(A,C) = \underline{P(A|C)}\ P(C)$

$P(A|B)$

$= P(A|C$

- [x] $A \perp\!\!\!\perp B$
- [ ] $A \perp\!\!\!\perp B \mid C$
- [x] $A \perp\!\!\!\perp C$
- [ ] $A \perp\!\!\!\perp C \mid B$
- [ ] $B \perp\!\!\!\perp C$
- [ ] $B \perp\!\!\!\perp C \mid A$
- [ ] No independence assumptions needed.

(ii) $\mathbf{P(A \mid B, C) = \frac{P(A)\ P(B|A)\ P(C|A)}{P(B|C)\ P(C)}}$ $P(C|A,B)$

- [ ] $A \perp\!\!\!\perp B$
- [ ] $A \perp\!\!\!\perp B \mid C$
- [ ] $A \perp\!\!\!\perp C$
- [ ] $A \perp\!\!\!\perp C \mid B$
- [ ] $B \perp\!\!\!\perp C$
- [x] $B \perp\!\!\!\perp C \mid A$
- [ ] No independence assumptions needed.

(iii) $\mathbf{P(A, B) = \sum_c P(A \mid B, c)\ P(B \mid c)\ P(c)}$

- [ ] $A \perp\!\!\!\perp B$
- [ ] $A \perp\!\!\!\perp B \mid C$
- [ ] $A \perp\!\!\!\perp C$
- [ ] $A \perp\!\!\!\perp C \mid B$
- [ ] $B \perp\!\!\!\perp C$
- [ ] $B \perp\!\!\!\perp C \mid A$
- [x] No independence assumptions needed.

$P(A,B,C,D) = P(B|A,C,D)\ P(A|C,D)\ P(C,D)$

(iv) $\mathbf{P(A,B \mid C,D) = P(A \mid C,D)\ P(B \mid A,C,D)}$

- [ ] $A \perp\!\!\!\perp B$
- [ ] $A \perp\!\!\!\perp B \mid C$
- [ ] $A \perp\!\!\!\perp B \mid D$
- [ ] $C \perp\!\!\!\perp D$
- [ ] $C \perp\!\!\!\perp D \mid A$
- [ ] $C \perp\!\!\!\perp D \mid B$
- [x] No independence assumptions needed.

(c) **(i)** Mark **all** expressions that are equal to $\mathbf{P(A \mid B)}$, **given no independence assumptions**.

- [ ] $\sum_c P(A \mid B, c)$
- [x] $\sum_c P(A, c \mid B)$
- [ ] $\frac{P(B|A)\ P(A|C)}{\sum_c P(B,c)}$
- [x] $\frac{\sum_c P(A,B,c)}{\sum_c P(B,c)}$
- [ ] $\frac{P(A,C|B)}{P(C|B)}$
- [ ] $\frac{P(A|C,B)\ P(C|A,B)}{P(C|B)}$
- [ ] None of the provided options.

(ii) Mark **all** expressions that are equal to $\mathbf{P(A,B,C)}$, given that $\mathbf{A \perp\!\!\!\perp B}$.

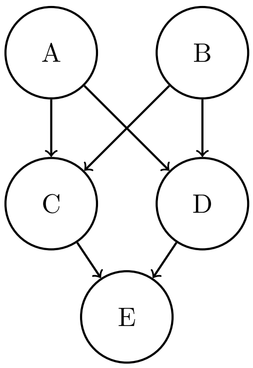
- [ ] $P(A \mid C)\ P(C \mid B)\ P(B)$
- [x] $P(A)\ P(B)\ P(C \mid A, B)$
- [ ] $P(C)\ P(A \mid C)\ P(B \mid C)$
- [ ] $P(A)\ P(C \mid A)\ P(B \mid C)$
- [x] $P(A)\ P(B \mid A)\ P(C \mid A, B)$
- [x] $P(A, C)\ P(B \mid A, C)$
- [ ] None of the provided options.

(iii) Mark **all** expressions that are equal to $\mathbf{P(A,B \mid C)}$, given that $\mathbf{A \perp\!\!\!\perp B \mid C}$.

- [x] $P(A \mid C)\ P(B \mid C)$
- [ ] $\frac{P(A)\ P(B|A)\ P(C|A,B)}{\sum_c P(A,B,c)}$
- [ ] $P(A \mid B)\ P(B \mid C)$
- [ ] $\frac{P(C)\ P(B|C)\ P(A|C)}{P(C|A,B)}$
- [ ] $\frac{\sum_c P(A,B,c)}{P(C)}$
- [x] $\frac{P(C,A|B)\ P(B)}{P(C)}$
- [ ] None of the provided options.

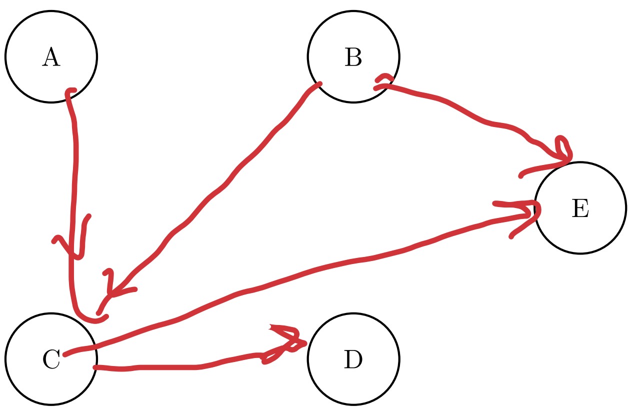
Bayes Nets and Joint Distributions

(a) Write down the joint probability distribution associated with the following Bayes Net. Express the answer as a product of terms representing individual conditional probabilities tables associated with this Bayes Net:



$P(A)\ P(B)\ P(C|A,B)$
$P(D|A,B)\ P(E|C,D)$

(b) Draw the Bayes net associated with the following joint distribution:
$P(A) \cdot P(B) \cdot P(C|A,B) \cdot P(D|C) \cdot P(E|B,C)$



(c) Do the following products of factors correspond to a valid joint distribution over the variables A, B, C, D? (Circle TRUEFALSE.)

(i) TRUE FALSE $P(A) \cdot P(B) \cdot P(C|A) \cdot P(C|B) \cdot P(D|C)$

(ii) TRUE FALSE $P(A) \cdot P(B|A) \cdot P(C) \cdot P(D|B,C)$

(iii) TRUE FALSE $P(A) \cdot P(B|A) \cdot P(C) \cdot P(C|A) \cdot P(D)$

(iv) TRUE FALSE $P(A|B) \cdot P(B|C) \cdot P(C|D) \cdot P(D|A)$ cycle

(d) What factor can be multiplied with the following factors to form a valid joint distribution? (Write "none" if the given set of factors can't be turned into a joint by the inclusion of exactly one more factor.)

(i) $P(A) \cdot P(B|A) \cdot P(C|A) \cdot P(E|B, C, D)$

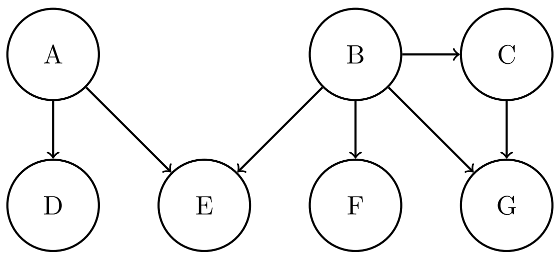
P (D) is missing

Draw BN

(ii) $P(D) \cdot P(B) \cdot P(C|D, B) \cdot P(E|C, D, A)$

P (A) is missing

(e) Answer the next questions based off of the Bayes Net below:
All variables have domains of {-1, 0, 1}



(i) Before eliminating any variables or including any evidence, how many entries does the factor at G have?

P (G | B, C)

$3^3 = 27$

(ii) Now we observe $e = 1$ and want to query $P(D|e = 1)$, and you get to pick the first variable to be eliminated.

- Which choice would create the **largest** factor f_1?

Eliminate B → highest connection

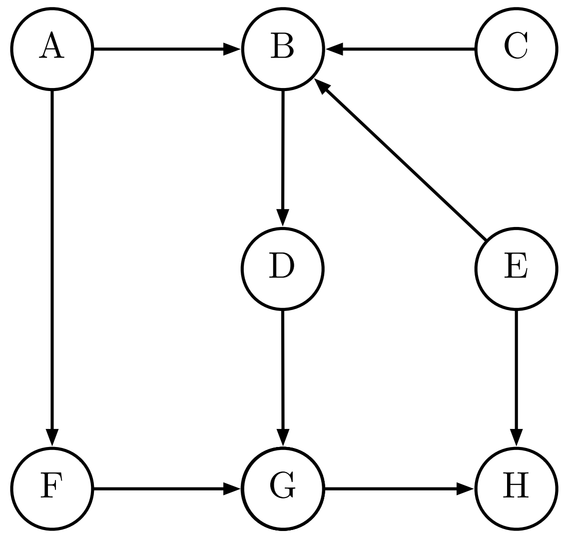
- Which choice would create the **smallest** factor f_1?

Eliminate A → smallest connection related to P(D | E)

Bayes' Nets Representation

(a) Graph Structure: Conditional Independence

Consider the Bayes' net given below.



Remember that $X \perp\!\!\!\perp Y$ reads as "X is independent of Y given nothing", and $X \perp\!\!\!\perp Y|\{Z, W\}$ reads as "X is independent of Y given Z and W."

For each expression, fill in the corresponding circle to indicate whether it is True or False.

(i) True False It is guaranteed that $A \perp\!\!\!\perp B$ line

(ii) True False It is guaranteed that $A \perp\!\!\!\perp C$ all inactive

(iii) True False It is guaranteed that $A \perp\!\!\!\perp D \mid \{B, H\}$ A F G D (active)
H

(iv) True False It is guaranteed that $A \perp\!\!\!\perp E|F$ all inactive

(v) True False It is guaranteed that $G \perp\!\!\!\perp E|B$ E B A F G (active)

(vi) True False It is guaranteed that $F \perp\!\!\!\perp C|D$ F A B C D (active)

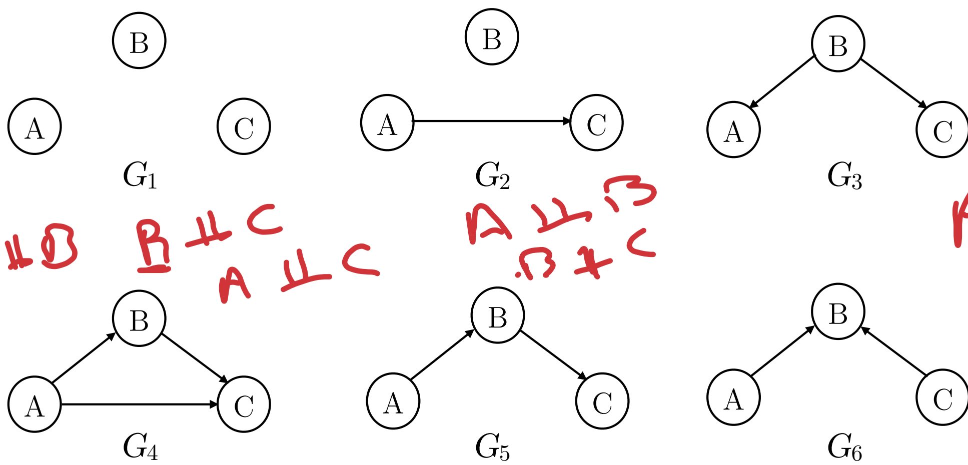
(vii) True False It is guaranteed that $E \perp\!\!\!\perp D|B$

(viii) True False It is guaranteeed that $C \perp\!\!\!\perp H|G$ A B E H
active

(b) Graph structure: Representational Power

Recall that any directed acyclic graph G has an associated family of probability distributions, which consists of all probability distributions that can be represented by a Bayes' net with structure G.

For the following questions, consider the following six directed acyclic graphs:



$A \not\perp\!\!\!\perp B$ $\quad B \not\perp\!\!\!\perp C$ $\quad A \perp\!\!\!\perp C$ $\quad A \perp\!\!\!\perp B$ $\quad B \not\perp\!\!\!\perp C$ $\quad A \not\perp\!\!\!\perp C | B$

$A \perp\!\!\!\perp C | B$ $\quad A \not\perp\!\!\!\perp C$

(i) Assume all we know about the joint distribution $P(A, B, C)$ is that it can be represented by the product $P(A|B, C)P(B|C)P(C)$. Mark each graph for which the associated family of probability distributions is guaranteed to include $P(A, B, C)$.

☐ G_1 ☐ G_2 ☐ G_3

☒ G_4 ☐ G_5 ☐ G_6

(ii) Now assume all we know about the joint distribution $P(A, B, C)$ is that it can be represented by the product $P(C|B)P(B|A)P(A)$. Mark each graph for which the associated family of probability distributions is guaranteed to include $P(A, B, C)$.

☐ G_1 ☐ G_2 ☒ G_3

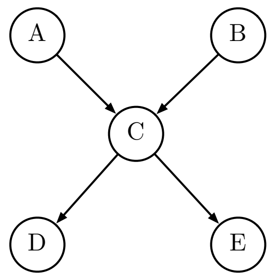
☒ G_4 ☒ G_5 ☐ G_6

(c) Marginalization and Conditioning

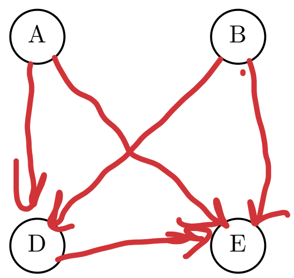
Consider a Bayes' net over the random variables A, B, C, D, E with the structure shown below, with full joint distribution $P(A, B, C, D, E)$.

The following three questions describe different, unrelated situations (your answers to one question should not influence your answer to other questions).

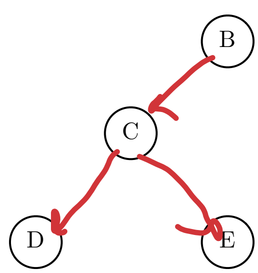
marginialization
make nieghbor
dependent



(i) Consider the marginal distribution $P(A, B, D, E) = \sum_c P(A, B, c, D, E)$, where C was eliminated. On the diagram below, draw the minimal number of arrows that results in a Bayes' net structure that is able to represent this marginal distribution. If no arrows are needed write "No arrows needed."

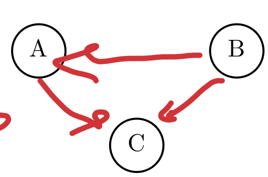


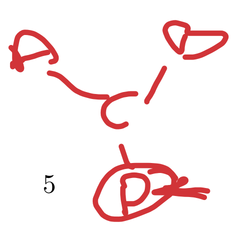
(ii) Assume we are given an observation: $A = a$. On the diagram below, draw the minimal number of arrows that results in a Bayes' net structure that is able to represent the conditional distribution $P(B, C, D, E \mid A = a)$. If no arrows are needed write "No arrows needed."



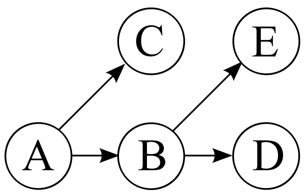
(iii) Assume we are given two observations: $D = d, E = e$. On the diagram below, draw the minimal number of arrows that results in a Bayes' net structure that is able to represent the conditional distribution $P(A, B, C \mid D = d, E = e)$. If no arrows are needed write "No arrows needed."

originally $A \perp\!\!\!\perp B$
but since D is
given
A now become
dependent





Bayes Nets: Variable Elimination



	$P(A)$
$+a$	0.25
$-a$	0.75

$P(B\|A)$	$+b$	$-b$
$+a$	0.5	0.5
$-a$	0.25	0.75

$P(C\|A)$	$+c$	$-c$
$+a$	0.2	0.8
$-a$	0.6	0.4

$P(D\|B)$	$+d$	$-d$
$+b$	0.6	0.4
$-b$	0.8	0.2

$P(E\|B)$	$+e$	$-e$
$+b$	0.25	0.75
$-b$	0.1	0.9

(a) Using the Bayes' Net and conditional probability tables above, calculate the following quantities:

(i) $P(+b|+a) =$ 0.5 Direct

(ii) $P(+a,+b) =$ $P(b|a)P(+a) = 0.5 \times 0.25$

(iii) $P(+a|+b) =$ $\frac{P(+a,+b)}{P(+b)} = \frac{0.5 \times 0.25}{0.25 \times 0.5 + 0.25 \times 0.75} = 0.4$

(b) Now we are going to consider variable elimination in the Bayes' Net above.

(i) Assume we have the evidence $+c$ and wish to calculate $P(E \mid +c)$. What factors do we have initially?

$P(A)\ P(B|A)\ P(+c|A)\ P(E|B)\ P(D|B)$

(ii) If we eliminate variable B, we create a new factor. What probability does that factor correspond to?

$P(D|B)\ P(E|B)\ P(B|A) \longrightarrow \underline{P(E,D|A)}$

(iii) What is the equation to calculate the factor we create when eliminating variable B?

$f(A,D,E) = \sum_B P(B|A)\ P(E|B)\ P(D|B)$

(iv) After eliminating variable B, what are the new set of factors? As in (ii), write the probabilities that the factors represent. For each factor, also provide its size.

$P(A)$ 2

$P(A,D,E)$ 8

$P(+c|A)$ 2

12

(v) Now assume we have the evidence $-c$ and are trying to calculate $P(A|-c)$. What is the most efficient elimination ordering? If more than one ordering is most efficient, provide any one of them.

small to large $E \to D \to B$ or $D \to E \to B$

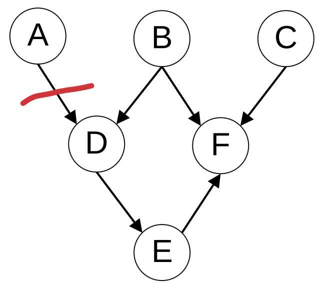
(vi) Once we have run variable elimination and have $f(A,-c)$ how do we calculate $P(+a \mid -c)$?

BN $\frac{f(+a,-c)}{P(-c)} = \frac{f(+a,-c)}{f(+a,-c) + f(-a,-c)}$

Bayes Nets: Independence

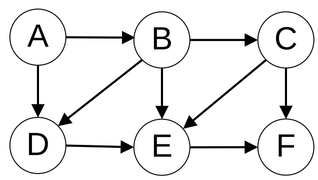
(a) For the following graphs, explicitly state the minimum size set of edges that must be removed such that the corresponding independence relations are guaranteed to be true.

Marked the removed edges with an 'X' on the graphs.



$A \perp\!\!\!\perp B|F$
$A \perp\!\!\!\perp F|D$
$B \perp\!\!\!\perp C$

AD



$A \perp\!\!\!\perp D|B$
$A \perp\!\!\!\perp F|C$
$C \perp\!\!\!\perp D|B$

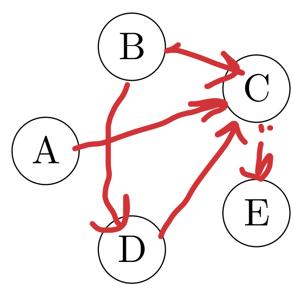
AD + EF or AB

(b) You're performing variable elimination over a Bayes Net with variables A, B, C, D, E. So far, you've finished joining over (but not summing out) C, when you realize you've lost the original Bayes Net!

Your current factors are $f(A)$, $f(B)$, $f(B,D)$, $f(A,B,C,D,E)$. Note: these are factors, NOT joint distributions. You don't know which variables are conditioned or unconditioned.

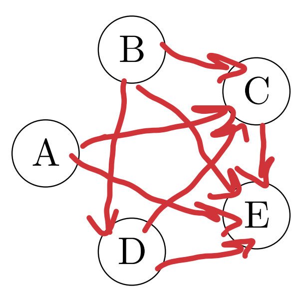
(i) What's the smallest number of edges that could have been in the original Bayes Net? Draw out one such Bayes Net below.

Number of edges = 5



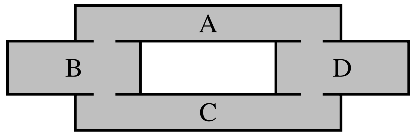
(ii) What's the largest number of edges that could have been in the original Bayes Net? Draw out one such Bayes Net below.

Number of edges = 8



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) 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) 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) 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						
$X^0 = \text{D}$	○ γ	○ $1 - \gamma$	○ $1 - 2\gamma$	○ 0.0	○ 1.0	○ 2γ	○ None of these
$X^1 = \text{B}$	○ γ	○ $1 - \gamma$	○ $1 - 2\gamma$	○ 0.0	○ 1.0	○ 2γ	○ None of these
$X^2 = \text{C}$	○ 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) 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
$X^0 = \text{D}$	○ $\gamma^2 - 2\gamma^3$ ○ $3 - 2\gamma$ ○ 0.0 ○ $\gamma - \gamma^2 + \gamma^3$ ◉ $1 - 3\gamma + 2\gamma^2$ ○ $2 - \gamma$ ○ $1 - 2\gamma + \gamma^2$ ○ None of these
$X^1 = \text{B}$	○ $\gamma^2 - 2\gamma^3$ ○ $3 - 2\gamma$ ◉ 0.0 ○ $\gamma - \gamma^2 + \gamma^3$ ○ $1 - 3\gamma + 2\gamma^2$ ○ $2 - \gamma$ ○ $1 - 2\gamma + \gamma^2$ ○ None of these
$X^2 = \text{C}$	○ $\gamma^2 - 2\gamma^3$ ○ $3 - 2\gamma$ ○ 0.0 ○ $\gamma - \gamma^2 + \gamma^3$ ○ $1 - 3\gamma + 2\gamma^2$ ○ $2 - \gamma$ ○ $1 - 2\gamma + \gamma^2$ ◉ None of these

$(1-\lambda)\times 1\times(1-2\lambda)$
$(1-$
0
$=(2\lambda)2\lambda$
$=(\lambda(1-\lambda))T$

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:

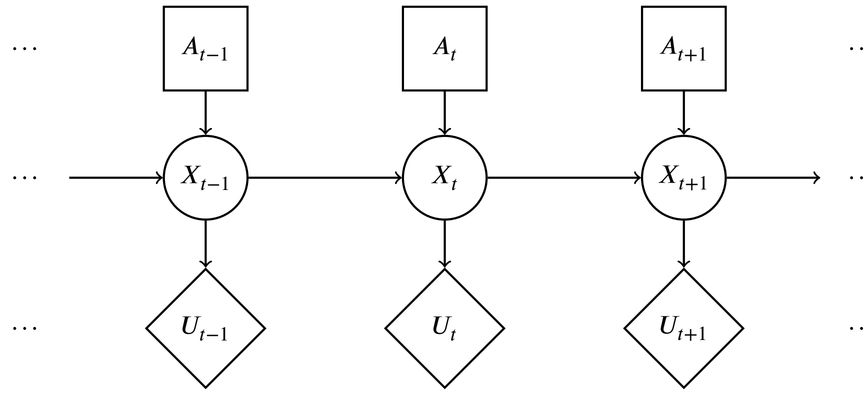
$$
\begin{aligned}
P(X_t \mid m^A_{0:t}, m^B_{0:t}, m^D_{0:t}) &\propto P(X_t, m^A_{0:t}, m^B_{0:t}, m^D_{0:t}) && (1)\\
&= \sum_{x_{t-1}} P(X_t, x_{t-1}, m^A_t, m^B_t, m^D_t, m^A_{0:t-1}, m^B_{0:t-1}, m^D_{0:t-1}) && (2)\\
&= \sum_{x_{t-1}} \underline{\qquad\qquad} P(X_t \mid x_{t-1}) P(x_{t-1}, m^A_{0:t-1},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

(e) 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

Planning ahead with HMMs

Pacman is tired of using HMMs to estimate the location of ghosts. He wants to use HMMs to plan what actions to take in order to maximize his utility. Pacman uses the HMM (drawn to the right) of length T to model the planning problem. In the HMM, $X_{1:T}$ is the sequence of hidden states of Pacman's world, $A_{1:T}$ are actions Pacman can take, and U_t is the utility Pacman receives at the particular hidden state X_t. Notice that there are no evidence variables, and utilities are not discounted.



(a) The belief at time t is defined as $B_t(X_t) = p(X_t|a_{1:t})$. The forward algorithm update has the following form:

$$B_t(X_t) = \underline{\quad\textbf{(i)}\quad}\ \underline{\quad\textbf{(ii)}\quad}\ B_{t-1}(x_{t-1}).$$

Complete the expression by choosing the option that fills in each blank.

(i) ○ $\max_{x_{t-1}}$ ○ $\sum_{x_{t-1}}$ ○ $\max_{x_t}$ ○ $\sum_{x_t}$ ○ 1

(ii) ○ $p(X_t|x_{t-1})$ ○ $p(X_t|x_{t-1})p(X_t|a_t)$ ○ $p(X_t)$ ○ $p(X_t|x_{t-1}, a_t)$ ○ 1

○ None of the above combinations is correct

(b) Pacman would like to take actions $A_{1:T}$ that maximizes the expected sum of utilities, which has the following form:

$$\text{MEU}_{1:T} = \underline{\ \textbf{(i)}\ }\ \underline{\ \textbf{(ii)}\ }\ \underline{\ \textbf{(iii)}\ }\ \underline{\ \textbf{(iv)}\ }\ \underline{\ \textbf{(v)}\ }$$

Complete the expression by choosing the option that fills in each blank.

(i) ○ $\max_{a_{1:T}}$ ○ $\max_{a_T}$ ○ $\sum_{a_{1:T}}$ ○ $\sum_{a_T}$ ○ 1

(ii) ○ $\max_t$ ○ $\prod_{t=1}^{T}$ ○ $\sum_{t=1}^{T}$ ○ $\min_t$ ○ 1

(iii) ○ $\sum_{x_t, a_t}$ ○ $\sum_{x_t}$ ○ $\sum_{a_t}$ ○ $\sum_{x_T}$ ○ 1

(iv) ○ $p(x_t|x_{t-1}, a_t)$ ○ $p(x_t)$ ○ $B_t(x_t)$ ○ $B_T(x_T)$ ○ 1

(v) ○ U_T ○ $\frac{1}{U_t}$ ○ $\frac{1}{U_T}$ ○ U_t ○ 1

○ None of the above combinations is correct

(c) A greedy ghost now offers to tell Pacman the values of some of the hidden states. Pacman needs your help to figure out if the ghost's information is useful. Assume that the transition function $p(x_t|x_{t-1}, a_t)$ is not deterministic. **With respect to the utility** U_t, mark all that can be True:

☐ $\text{VPI}(X_{t-1}|X_{t-2}) > 0$ ☐ $\text{VPI}(X_{t-2}|X_{t-1}) > 0$ ☐ $\text{VPI}(X_{t-1}|X_{t-2}) = 0$ ☐ $\text{VPI}(X_{t-2}|X_{t-1}) = 0$ ☐ None of the above

(d) Pacman notices that calculating the beliefs under this model is very slow using exact inference. He therefore decides to try out various particle filter methods to speed up inference. Order the following methods by how accurate their estimate of $B_T(X_T)$ is? If different methods give an equivalently accurate estimate, mark them as the same number.

	Most accurate			Least accurate
Exact inference	○ 1	○ 2	○ 3	○ 4
Particle filtering with no resampling	○ 1	○ 2	○ 3	○ 4
Particle filtering with resampling before every time elapse	○ 1	○ 2	○ 3	○ 4
Particle filtering with resampling before every other time elapse	○ 1	○ 2	○ 3	○ 4