CS188: Exam Practice Session 7 Solutions

Q1. Probabilities

(a) Fill in the circles of **all** expressions that are equal to **1**,
given no independence assumptions:

- ● $\sum_a P(A = a \mid B)$
- ○ $\sum_b P(A \mid B = b)$
- ● $\sum_a \sum_b P(A = a, B = b)$
- ○ $\sum_a \sum_b P(A = a \mid B = b)$
- ● $\sum_a \sum_b P(A = a)\ P(B = b)$
- ○ $\sum_a P(A = a)\ P(B = b)$
- ○ None of the above.

Probability distributions, including conditional distributions, sum to one. We are testing this axiom when applied to multivariate distributions and conditionals.

(b) Fill in the circles of **all** expressions that are equal to $\mathbf{P(A, B, C)}$,
given no independence assumptions:

- ● $P(A \mid B, C)\ P(B \mid C)\ P(C)$
- ○ $P(C \mid A, B)\ P(A)\ P(B)$
- ● $P(A, B \mid C)\ P(C)$
- ● $P(C \mid A, B)\ P(A, B)$
- ○ $P(A \mid B)\ P(B \mid C)\ P(C)$
- ○ $P(A \mid B, C)\ P(B \mid A, C)\ P(C \mid A, B)$
- ○ None of the above.

We are testing the chain rule when applied to more than two variables.

(c) Fill in the circles of **all** expressions that are equal to $\mathbf{P(A \mid B, C)}$,
given no independence assumptions:

- ● $\frac{P(A,B,C)}{\sum_a P(A=a,B,C)}$
- ● $\frac{P(B,C|A)\ P(A)}{P(B,C)}$
- ● $\frac{P(B|A,C)\ P(A|C)}{P(B|C)}$
- ○ $\frac{P(B|A,C)\ P(A|C)}{P(B,C)}$
- ○ $\frac{P(B|A,C)\ P(C|A,B)}{P(B,C)}$
- ○ $\frac{P(A,B|C)}{P(B|A,C)}$
- ○ None of the above.

This is Bayes' rule applied to distributions over multiple variables. $P(A \mid B, C) = P(A, B, C)/P(B, C)$

(d) Fill in the circles of **all** expressions that are equal to $\mathbf{P(A \mid B)}$,
given that $\mathbf{A \perp\!\!\!\perp B \mid C}$:

- ○ $\frac{P(A|C)\ P(B|C)}{P(B)}$
- ○ $\frac{P(A|C)\ P(B|C)}{P(B|C)}$
- ● $\frac{\sum_c P(A|C=c)\ P(B|C=c)\ P(C=c)}{\sum_{c'} P(B|C=c')\ P(C=c')}$
- ○ $\frac{P(A|B,C)}{P(A|C)}$
- ● $\frac{\sum_c P(B|A,C=c)\ P(A,C=c)}{P(B)}$
- ○ $\frac{\sum_c P(A,C=c)\ P(B|C=c)}{\sum_{c'} P(A,B,C=c')}$
- ○ None of the above.

Apply Bayes' rule to get $P(A \mid B) = P(A, B)/P(B) = \sum_c P(A, B \mid C = c)P(C = c)/P(B)$ and conditional independence $P(A, B \mid C) = P(A \mid C)\ P(B \mid C)$

(e) Fill in the circles of **all** expressions that are equal to $\mathbf{P(A, B, C)}$, **given that** $\mathbf{A \perp\!\!\!\perp B \mid C}$ **and** $\mathbf{A \perp\!\!\!\perp C}$:

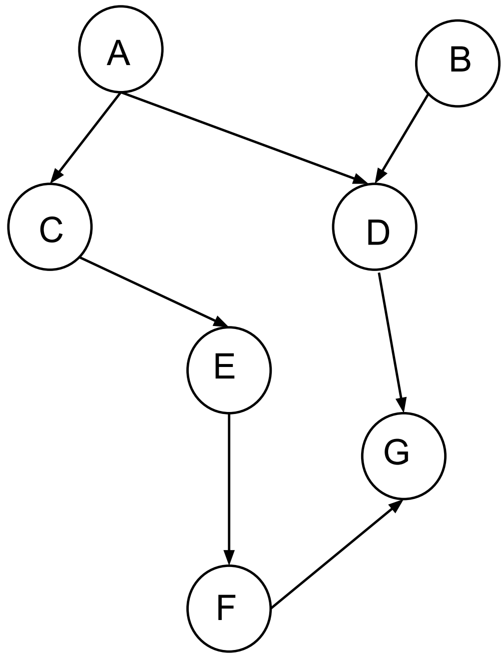
- ○ $P(A)\ P(B)\ P(C)$
- ● $P(A)\ P(B, C)$
- ● $P(A \mid B)\ P(B \mid C)\ P(C)$
- ○ $P(A \mid B, C)\ P(B \mid A, C)\ P(C \mid A, B)$
- ● $P(A \mid C)\ P(B \mid C)\ P(C)$
- ○ $P(A \mid C)\ P(B \mid C)$
- ○ None of the above.

If $A \perp\!\!\!\perp B \mid C$ and $A \perp\!\!\!\perp C$ it can be proven that $A \perp\!\!\!\perp B$ but not that $B \perp\!\!\!\perp C$. Here is the proof:

$$\begin{aligned}
P(A, B) &= \sum_c P(A, B \mid C = c) P(C = c) \\
&= \sum_c P(A \mid C = c) P(B \mid C = c) P(C = c) \\
&= P(A) \sum_c P(B \mid C = c) P(C = c) \\
&= P(A) \sum_c P(B, C = c) \\
&= P(A)\ P(B)
\end{aligned}$$

Q2. Bayes Nets: Independence

Consider a Bayes Net with the following graph:



Which of the following are guaranteed to be true without making any additional conditional independence assumptions, other than those implied by the graph? (Mark all true statements)

- ● $P(A \mid C, E) = P(A \mid C)$
- ○ $P(A, E \mid G) = P(A \mid G) * P(E \mid G)$
- ● $P(A \mid B = b) = P(A)$
- ○ $P(A \mid B, G) = P(A \mid G)$
- ○ $P(E, G \mid D) = P(E \mid D) * P(G \mid D)$
- ● $P(A, B \mid F) = P(A \mid F) * P(B \mid F)$

This question deals with (conditional) independence of a Bayes Net.
Option 1: $A \perp\!\!\!\perp E \mid C$, since with C observed no path between A and C is active.
Option 2: there's no conditional independence between A and E given G, since A-C-E is active.
Option 3: $A \perp\!\!\!\perp B$, since no path between A and B is active.
Option 4: there's no conditional independence between A and B given G, since the path A, B, G is an active.
Option 5: there's no conditional independence between E and G, since E-F-G is active.
Option 6: $A \perp\!\!\!\perp B \mid F$, since A, B, F is active.

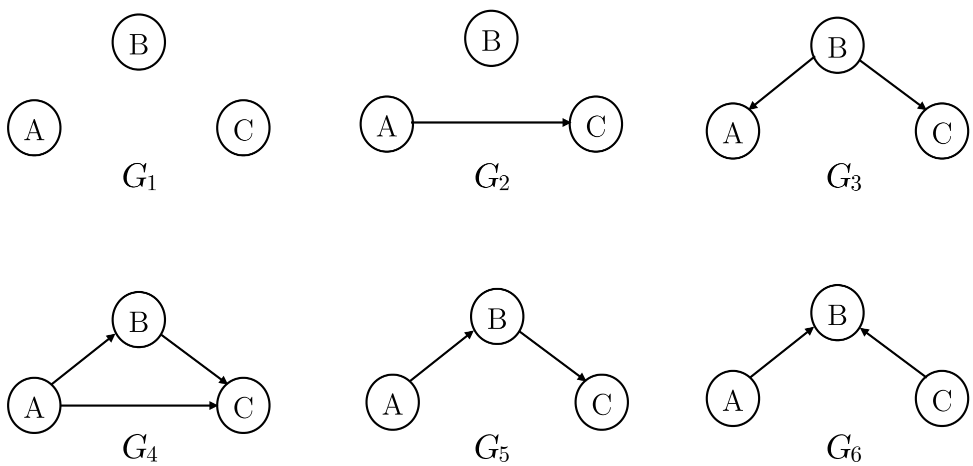
Q3. Bayes Nets: Representation

(a) 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:

In general, the absence of an edge implies independence but the presence of an edge does not guarantee dependence. For a Bayes' net to represent a joint distribution, it can only make a subset of the conditional independence assumptions given by the joint. If a Bayes' net makes more independence assumptions than the joint, its family of distributions is not guaranteed to include the joint distribution because the Bayes' net family is constrained by more independence relationships. For instance G_1 can only represent the completely independent joint $P(A, B, C) = P(A)P(B)P(C)$.



(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

G_4 is fully connected, and is therefore able to represent any joint distribution. The others cannot represent $P(A|B, C)P(B|C)P(C)$ because they make more independence assumptions, which you can verify. For example, G_3 assumes $A \perp\!\!\!\perp C|B$ but this is not given by the joint.

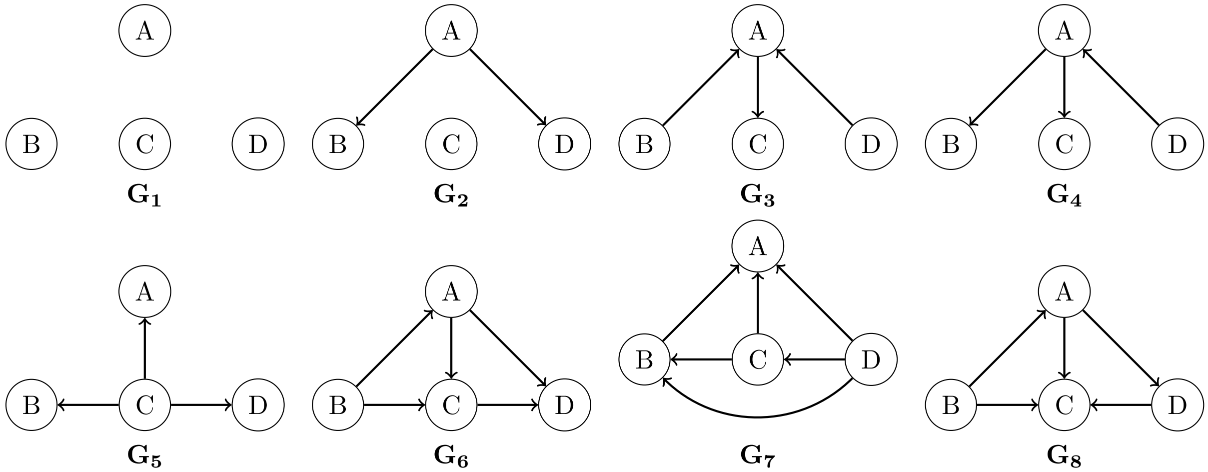
(ii) Now assume all we know about the joint distribution $P(A, B, C)$ is that it can be represented by the product $P(C|B)P(B|A)P(A)$. Mark each graph for which the associated family of probability distributions is guaranteed to include $P(A, B, C)$.

○ G_1 ○ G_2 ● G_3

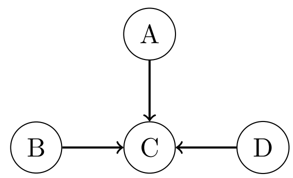
● G_4 ● G_5 ○ G_6

G_1 assumes all variables are independent, G_2 B is independent of the others, and G_6 assumes $A \perp\!\!\!\perp C$.

(b) For the following questions, consider the following eight directed acyclic graphs:



(i) Consider the Bayes' Net $\mathbf{B_1}$ below, and fill in **all the circles** (or select *None of the above*) corresponding to the Bayes' Nets $\mathbf{G_1}$ through $\mathbf{G_8}$ that are able to represent **at least one distribution** that $\mathbf{B_1}$ is able to represent.



$\mathbf{B_1}$

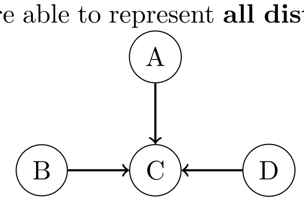
● $\mathbf{G_1}$ ● $\mathbf{G_2}$ ● $\mathbf{G_3}$ ● $\mathbf{G_4}$

● $\mathbf{G_5}$ ● $\mathbf{G_6}$ ● $\mathbf{G_7}$ ● $\mathbf{G_8}$

○ None of the above.

Consider the fully independent joint $P(A, B, C, D) = P(A)P(B)P(C)P(D)$ with uniform distribution (that is, every table entry has probability $\frac{1}{16}$. This can be represented by any Bayes' net. Pick any conditional independence assumption and verify that it is satisfied with by this distribution.

(ii) Consider the Bayes' Net $\mathbf{B_1}$ below, and fill in **all the circles** (or select *None of the above*) corresponding to the Bayes' Nets $\mathbf{G_1}$ through $\mathbf{G_8}$ that are able to represent **all distributions** that $\mathbf{B_1}$ is able to represent.



$\mathbf{B_1}$

○ $\mathbf{G_1}$ ○ $\mathbf{G_2}$ ○ $\mathbf{G_3}$ ○ $\mathbf{G_4}$

○ $\mathbf{G_5}$ ○ $\mathbf{G_6}$ ● $\mathbf{G_7}$ ● $\mathbf{G_8}$

○ None of the above.

To represent all of the distributions of B_1, a Bayes' net must not make further independence assumptions. A family of distributions that makes only a subset of the independences assumptions of B_1 can represent all of the same distributions as B_1.

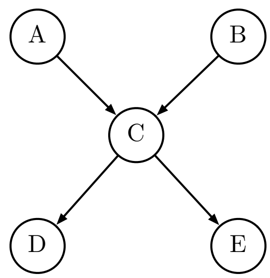
(c) Marginalization and Conditioning

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

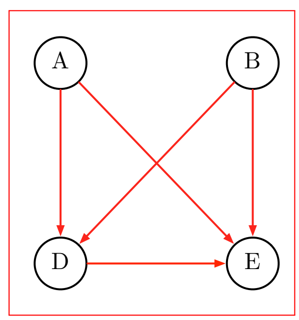
The following three questions describe different, unrelated situations (your answers to one question should not influence your answer to other questions).

Marginalization renders the neighbors of the marginalized out variable dependent.

Conditioning fixes the observed variables and renders their ancestors dependent according to the rules of d-separation.



(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."



The high level overview for these types of problems is that the resultant graph must be able to encode the same conditional independence assumptions from the initial Bayes' net we have. For example, let's look at the BN above. We see the following independence assumptions:

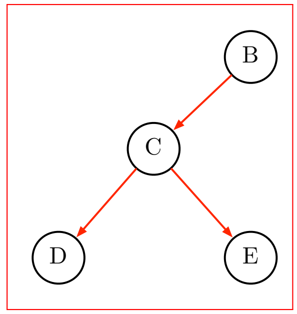
- $A \perp B$
- $A \perp D|C$
- $B \perp D|C$
- $A \perp E|C$
- $B \perp E|C$
- $D \perp E|C$

When we marginalize out C, we remove C from the graph. The conditional independence assumptions involving C no longer matter, so we just need to preserve:

$$A \perp B$$

To do this, we cannot have an edge between A and B. A and B must also be D-separated in the resultant BN, which it is in the solution. Every other edge is fair game because we don't make any other conditional independence assumptions.

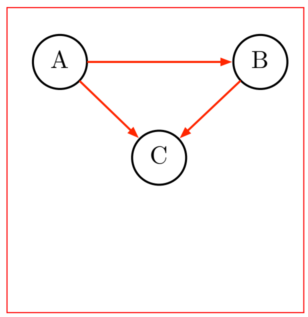
If you think about it, having $E \rightarrow D$ or $D \rightarrow E$ will fit the requirement above (it's also a valid point to say that the BN is symmetrical so the direction should not matter.) However, the arrow between A and D matters because we want ADB to be a common effect triple, which is an inactive triple if the middle node is unobserved, hence preserving the $A \perp B$ requirement.



(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."

Observing A fixes its value and removes it from the Bayes' net. By d-separation no further dependence is introduced.

(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." Observing D and E



makes an active path to their parent C, which in turn activates the common effect triple, and renders A and B dependent.