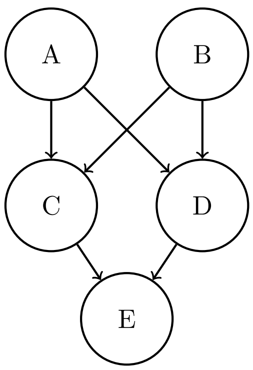
CS 188
Fall 2019

Exam Prep 6 Solutions

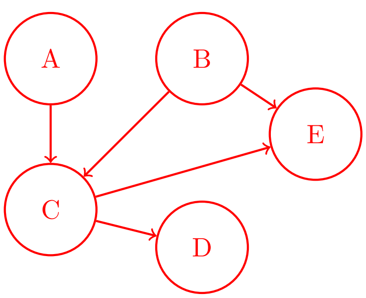
Q1. Bayes Nets and Joint Distributions

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

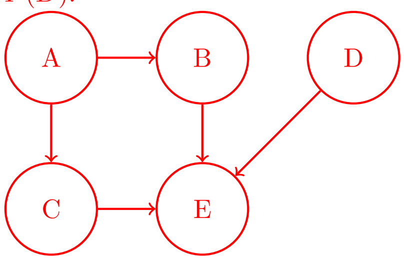
(iii) TRUE **FALSE** $P(A) \cdot P(B|A) \cdot P(C) \cdot P(C|A) \cdot P(D)$

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

(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.)

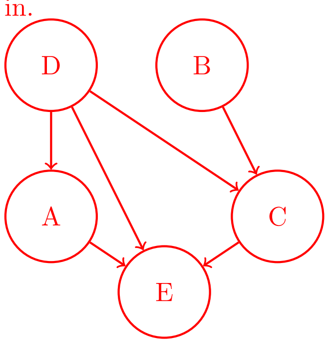
(i) $P(A) \cdot P(B|A) \cdot P(C|A) \cdot P(E|B, C, D)$

P(D) is missing. D could also be conditioned on A,B, and/or C without creating a cycle (e.g. $P(D|A, B, C)$). Here is an example bayes net that would represent the distribution after adding in $P(D)$:

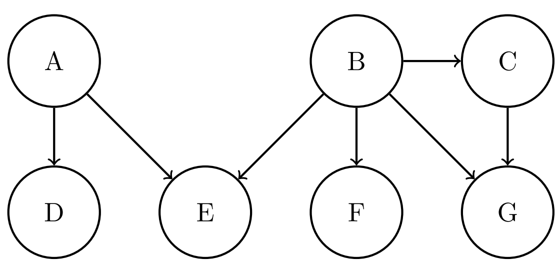


(ii) $P(D) \cdot P(B) \cdot P(C|D, B) \cdot P(E|C, D, A)$

P(A) is missing to form a valid joint distributions. A could also be conditioned on B, C, and/or D (e.g. $P(A|B, C, D)$. Here is a bayes net that would represent the distribution is $P(A|D)$ was added in.



(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?

The factor is $P(G|B, C)$, so that gives $3^3 = 27$ entries.

(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.

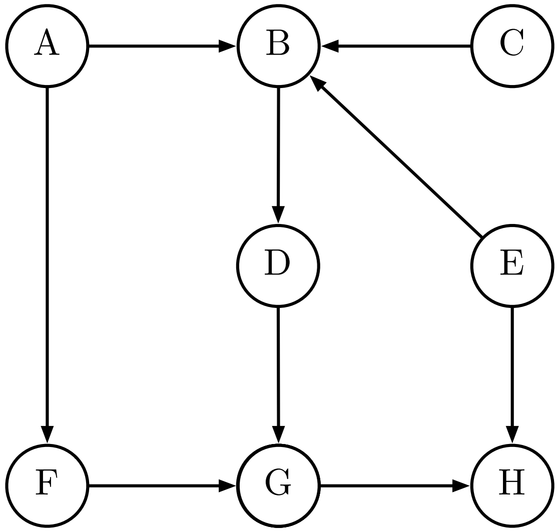
- Which choice would create the **largest** factor f_1?
  Eliminating B first would give the largest f_1:, $f_1(A, F, G, C, e) = \sum_{B=b} P(b)P(e|A, b)P(F|b)P(G|b, C)P(C|b)$. This factor has 3^4 entries.
- Which choice would create the **smallest** factor f_1?
  Eliminating A or eliminating F first would give smallest factors of 3 entries: either $f1(D, e) =$

$\sum_a P(D|a)P(e|a)P(a)$ of $f1(B) = \sum_f P(f|B)$. Eliminating D is not correct because D is the query variable.

Q2. 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 $A \perp\!\!\!\perp B$

The edge between A and B implies that independency is not guaranteed.

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

No active paths.

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

An active path: $A \rightarrow F \rightarrow G$ (descendant H observed) $\leftarrow D$.

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

No active paths.

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

An active path: $G \leftarrow F \leftarrow A \rightarrow B$ (observed) $\leftarrow E$.

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

An active path: $F \leftarrow A \rightarrow B$ (descendent D observed) $\leftarrow C$.

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

No active paths.

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

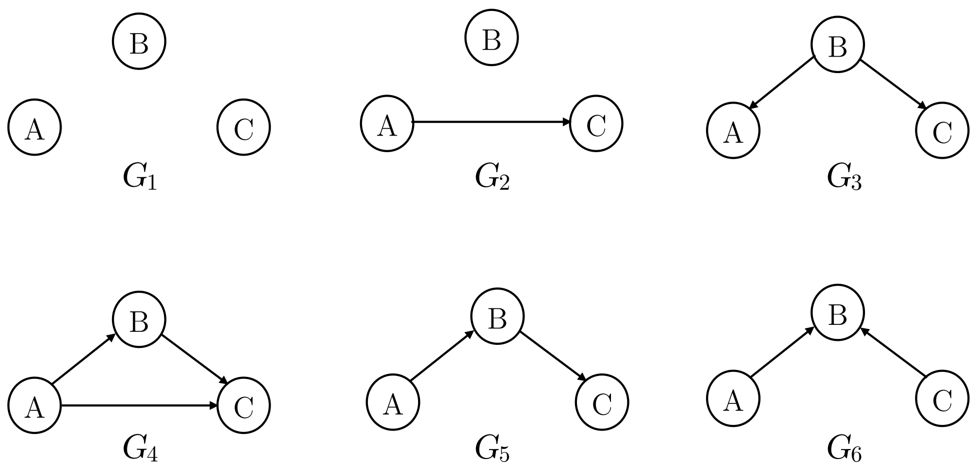
An active path: $C \to B$ (descendent G observed) $\leftarrow E \to H$.

(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:

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 $A \perp\!\!\!\perp C$.

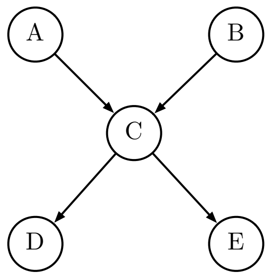
(c) Marginalization and Conditioning

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

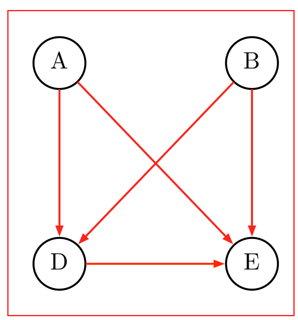
The following three questions describe different, unrelated situations (your answers to one question should not influence your answer to other questions).

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:

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

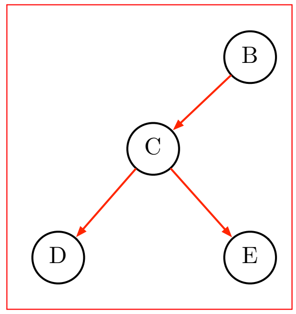
When we marginalize out C, we remove C from the graph. The conditional independence assumptions involving C no longer matter, so we just need to preserve:

$$A \perp B$$

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.

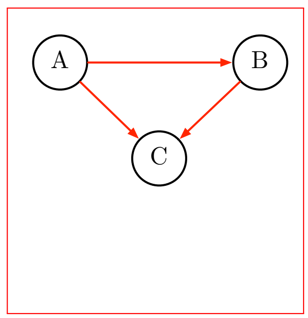
If you think about it, having $E \to D$ or $D \to E$ will fit the requirement above (it's also a valid point to say that the BN is symmetrical so the direction should not matter.) However, the arrow between A and D matters because we want ADB to be a common effect triple, which is an inactive triple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