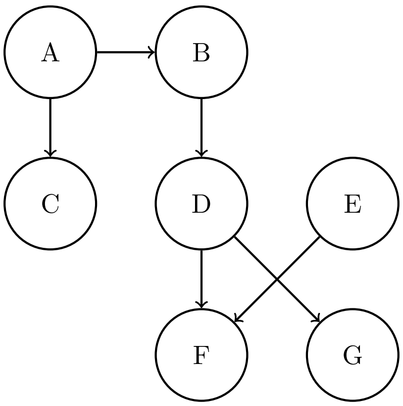
CS 188 Summer 2023 — Discussion 3A Solutions

1 Bayes Nets: Representation

Parts (a), (b), and (c) pertain to the following Bayes' Net.



(a) Express the joint probability distribution as a product of terms from the Bayes Nets CPTs.

$P(A)P(C|A)P(B|A)P(D|B)P(E)P(F|D,E)P(G|D)$

(b) Assume each node can take on 4 values. How many entries do the factors at A, D, and F have?

A: 4 D: 4^2 F: 4^3

(c) Mark all that are guaranteed to be true:

- ☐ $B \perp\!\!\!\perp C$
- ☐ $A \perp\!\!\!\perp F$
- ☐ $D \perp\!\!\!\perp E|F$
- ■ $E \perp\!\!\!\perp A|D$
- ■ $F \perp\!\!\!\perp G|D$
- ■ $B \perp\!\!\!\perp F|D$
- ☐ $C \perp\!\!\!\perp G$
- ■ $D \perp\!\!\!\perp E$

Parts (d) and (e) pertain to the following CPTs.

A	$P(A)$
+a	0.8
–a	0.2

A	B	$P(B\|A)$
+a	+b	0.9
+a	–b	0.1
–a	+b	0.6
–a	–b	0.4

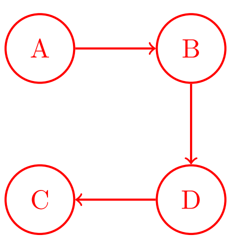
B	C	$P(C\|B)$
+b	+c	0.8
+b	–c	0.2
–b	+c	0.8
–b	–c	0.2

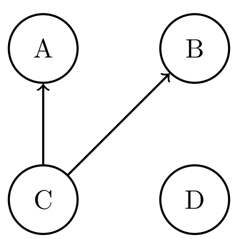
C	D	$P(D\|C)$
+c	+d	0.25
+c	–d	0.75
–c	+d	0.5
–c	–d	0.5

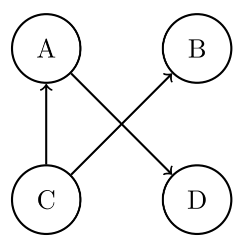
(d) State all non-conditional independence assumptions that are implied by the probability distribution tables.

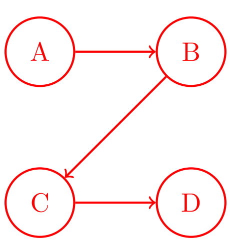
From the tables, we have $A \not\perp\!\!\!\perp B$ and $C \not\perp\!\!\!\perp D$, and $B \perp\!\!\!\perp C$.

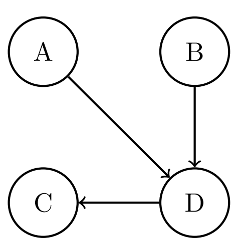
(e) Circle all the Bayes net(s) that can represent a distribution that is consistent with the tables given.

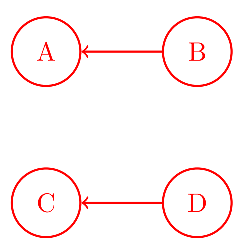












The question asks for Bayes Nets that **can** represent the distribution in the tables. So, in the nets we circle, the only requirement must be that A and B must not be independent, and C and D must not be independent.

The top left, bottom left, and bottom right nets have arrows between the A-B nodes and the C-D nodes, so we can circle those.

The top middle net has C and D as independent (D is not connected to anything), so we cannot circle it. The bottom middle net has A and B as independent (common effect), so we cannot circle it.

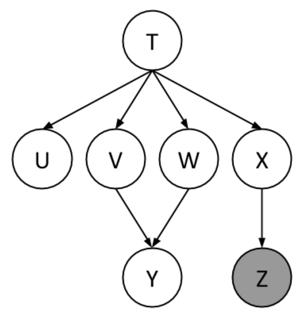
The top right net seems like it could represent the distribution, because D-separation finds that: A and B are not guaranteed to be independent (common cause), and C and D are not guaranteed to be independent (causal chain). However, since $B \perp\!\!\!\perp C$, the arrow between B and C is vacuous, thus A and B cannot actually be dependent.

2 Variable Elimination

Using the Bayes Net shown below, we want to compute $P(Y \mid +z)$. All variables have **binary domains**. We run variable elimination, with the following variable elimination ordering: X, T, U, V, W.

After inserting evidence, we have the following factors to start out with:

$$P(T), P(U|T), P(V|T), P(W|T), P(X|T), P(Y|V,W), P(+z|X)$$



(a) When eliminating X we generate a new factor f_1 as follows,

$$f_1(+z|T) = \sum_x P(x|T)P(+z|x)$$

which leaves us with the factors:

$$P(T), P(U|T), P(V|T), P(W|T), P(Y|V,W), f_1(+z|T)$$

(b) When eliminating T we generate a new factor f_2 as follows, which leaves us with the factors:

$$f_2(U,V,W,+z) = \sum_t P(t)P(U|t)P(V|t)P(W|t)f_1(+z|t) \qquad P(Y|V,W), f_2(U,V,W,+z)$$

(c) When eliminating U we generate a new factor f_3 as follows, which leaves us with the factors:

$$f_3(V,W,+z) = \sum_u f_2(u,V,W,+z) \qquad P(Y|V,W), f_3(V,W,+z)$$

Note that U could have just been deleted from the original graph, because $\sum_u P(U|t) = 1$. We can see this in the graph: we can remove any leaf node that is not a query variable or an evidence variable.

(d) When eliminating V we generate a new factor f_4 as follows, which leaves us with the factors:

$$f_4(W,Y,+z) = \sum_v f_3(v,W,+z)P(Y|v,W) \qquad f_4(W,Y,+z)$$

(e) When eliminating W we generate a new factor f_5 as follows, which leaves us with the factors:

$$f_5(Y, +z) = \sum_w f_4(w, Y, +z) \qquad f_5(Y, +z)$$

(f) How would you obtain $P(Y \mid +z)$ from the factors left above: Simply renormalize $f_5(Y, +z)$ to obtain $P(Y \mid +z)$. Concretely,

$$P(y \mid +z) = \frac{f_5(y, +z)}{\sum_{y'} f_5(y', +z)}$$

(g) What is the size of the largest factor that gets generated during the above process?
$f_2(U, V, W, +z)$. This contains 3 unconditioned variables, so it will have $2^3 = 8$ entries (U, V, W are binary variables, and we only need to store the entries for $+z$ for each possible setting of these variables).

(h) Does there exist a better elimination ordering (one which generates smaller largest factors)?
Yes. One such ordering is X, U, T, V, W. All factors generated with this ordering contain at most 2 unconditioned variables, so the tables will have at most $2^2 = 4$ entries (as all variables are binary).