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: 4² F: 4³

(c) Mark all that are guaranteed to be true:
   - B ⊥⊥ C  
   - A ⊥⊥ F  
   - D ⊥⊥ E|F  
   - E ⊥⊥ A|D  
   - F ⊥⊥ G|D  
   - B ⊥⊥ F|D  
   - C ⊥⊥ G  
   - D ⊥⊥ E

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

|   | A   | P(A) | B   | P(B|A) | C   | P(C|B) | D   | P(D|C) |
|---|-----|------|-----|--------|-----|--------|-----|--------|
| +a| 0.8 | +a   | b   | 0.9    | +b  | +c     | 0.8 | +c     |
| -a| 0.2 | +a   | -b  | 0.1    | +b  | -c     | 0.2 | +c     |
| -a| 0.2 | -a   | b   | 0.6    | -b  | +c     | 0.8 | -c     |
|   |     | -a   | -b  | 0.4    | -b  | -c     | 0.2 | -c     |

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

   From the tables, we have A ⊥⊥ B and C ⊥⊥ D, and B ⊥⊥ 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:

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

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

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

which leaves us with the factors:

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

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

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

$$P(Y\mid V, W), f_2(U, V, W, +z)$$

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

$$f_3(V, W, +z) = \sum_u f_2(u, V, W, +z)$$

$$P(Y\mid V, W), f_3(V, W, +z)$$

Note that $U$ could have just been deleted from the original graph, because $\sum_u P(U\mid t) = 1$. 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:

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

$$f_4(W, Y, +z)$$
(e) When eliminating $W$ we generate a new factor $f_5$ as follows, which leaves us with the factors:

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

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

$$P(y | +z) = \frac{f_5(y, +z)}{\sum_y f_5(y', +z)}$$

(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).