Homework 7 Written

July 8th, 2020 at 11:59pm

1 Bayes' Net: Representation

1. Assume we know that a joint distribution d_1(over A,B,C) can be represented by Bayes' net $\mathbf{B_1}$. Mark all of the following Bayes' nets that are guaranteed to be able to represent d_1.

   ☐ $\mathbf{G_1}$ ☐ $\mathbf{G_2}$ ☐ $\mathbf{G_3}$ ☐ $\mathbf{G_4}$ ☐ $\mathbf{G_5}$

   ☐ $\mathbf{G_6}$ ☐ $\mathbf{G_7}$ ☐ $\mathbf{G_8}$ ☐ $\mathbf{G_9}$ ☐ $\mathbf{G_{10}}$

   ☐ None of the above.

2. Assume we know that a joint distribution d_2(over A,B,C) can be represented by Bayes' net $\mathbf{B_2}$. Mark all of the following Bayes' nets that are guaranteed to be able to represent d_2.

   ☐ $\mathbf{G_1}$ ☐ $\mathbf{G_2}$ ☐ $\mathbf{G_3}$ ☐ $\mathbf{G_4}$ ☐ $\mathbf{G_5}$

   ☐ $\mathbf{G_6}$ ☐ $\mathbf{G_7}$ ☐ $\mathbf{G_8}$ ☐ $\mathbf{G_9}$ ☐ $\mathbf{G_{10}}$

   ☐ None of the above.

3. Assume we know that a joint distribution d_3(over A,B,C) can be represented by Bayes' net $\mathbf{B_3}$. Mark all of the following Bayes' nets that are guaranteed to be able to represent d_3.

   ☐ $\mathbf{G_1}$ ☐ $\mathbf{G_2}$ ☐ $\mathbf{G_3}$ ☐ $\mathbf{G_4}$ ☐ $\mathbf{G_5}$

   ☐ $\mathbf{G_6}$ ☐ $\mathbf{G_7}$ ☐ $\mathbf{G_8}$ ☐ $\mathbf{G_9}$ ☐ $\mathbf{G_{10}}$

   ☐ None of the above.

4. Assume we know that a joint distribution d_4(over A,B,C) can be represented by Bayes' net $\mathbf{B_1 B_2}$ and $\mathbf{B_3}$. Mark all of the following Bayes' nets that are guaranteed to be able to represent d_4.

   ☐ $\mathbf{G_1}$ ☐ $\mathbf{G_2}$ ☐ $\mathbf{G_3}$ ☐ $\mathbf{G_4}$ ☐ $\mathbf{G_5}$

   ☐ $\mathbf{G_6}$ ☐ $\mathbf{G_7}$ ☐ $\mathbf{G_8}$ ☐ $\mathbf{G_9}$ ☐ $\mathbf{G_{10}}$

   ☐ None of the above.

2 Variable Elimination

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

Solution:

$$P(A), P(B \mid A), P(+c), P(D \mid A, B, +c), P(E \mid D), P(F \mid D), P(G \mid +c, F)$$

When eliminating B we generate a new factor f_1 as follows:

Solution:

$$f_1(A, +c, D) = \sum_b P(b \mid A) P(D \mid A, b, +c)$$

This leaves us with the factors:

Solution:

$$P(A), P(+c), P(E \mid D), P(F \mid D), P(G \mid +c, F), f_1(A, +c, D)$$

When eliminating D we generate a new factor f_2 as follows:

Solution:

$$f_2(A, +c, E, F) = \sum_d P(E \mid d) P(F \mid d) f_1(A, +c, d)$$

This leaves us with the factors:

Solution:

$$P(A), P(+c), P(G \mid +c, F), f_2(A, +c, E, F)$$

When eliminating G we generate a new factor f_3 as follows:

Solution:

$$f_3(+c, F) = \sum_g P(g \mid +c, F)$$

This leaves us with the factors:

Solution:

$$P(A), P(+c), f_2(A, +c, E, F), f_3(+c, F)$$

When eliminating F we generate a new factor f_4 as follows:

Solution:

$$f_4(A, +c, E) = \sum_f f_2(A, +c, E, f) f_3(+c, f)$$

This leaves us with the factors:

Solution:

$$P(A), P(+c), f_4(A, +c, E)$$

(b) Write a formula to compute $P(A, E| + c)$ from the remaining factors.

Solution:

$$P(A, E| + c) = \frac{P(A)P(+c)f_4(A, +c, E)}{\sum_{a,e} P(a)P(+c)f_4(a, +c, e)}$$

(c) Among f_1, f_2, f_3, f_4, which is the largest factor generated, and how large is it? Assume all variables have binary domains and measure the size of each factor by the number of rows in the table that would represent the factor.

Solution: $f_2(A, +c, E, F)$ is the largest factor generated and it has 8 rows.

(d) Find a variable elimination ordering for the same query, i.e., for $P(A, E| + c)$, for which the maximum size factor generated along the way is smallest. Hint: the maximum size factor generated in your solution should have only 2 variables, for a size of $2^2 = 4$ table. Fill in the variable elimination ordering and the factors generated into the table below.

Variable Eliminated	Factor Generated
B	$f_1(A, +c, D)$
G	$f_2(+c, F)$
F	$f_3(+c, D)$
D	$f_4(A, +c, E)$

For example, in the naive ordering we used earlier, the first row in this table would have had the following two entries: B, $f_1(A, +c, D)$.