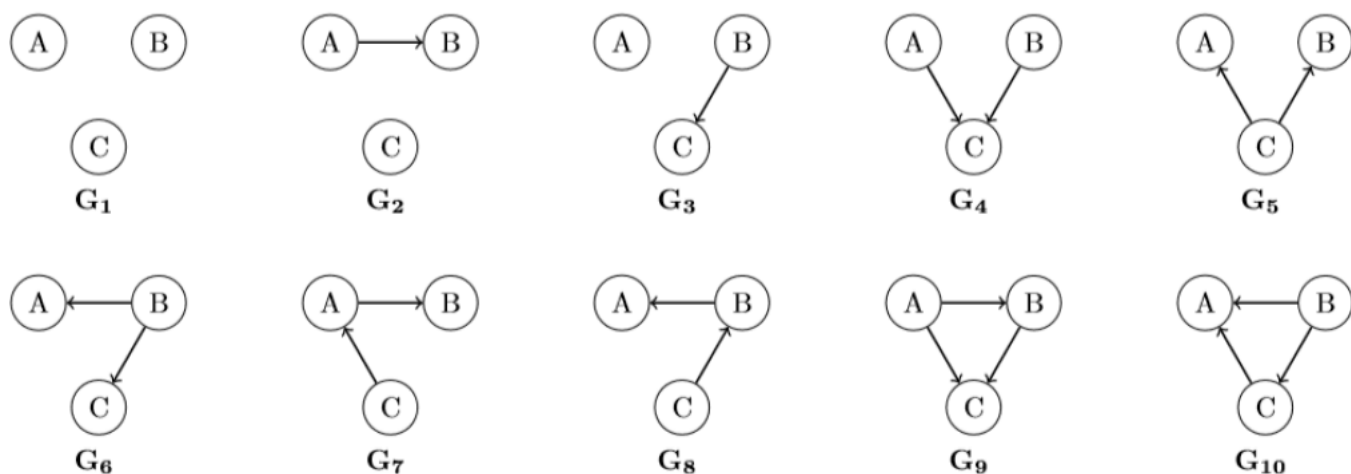


## Homework 6 Written

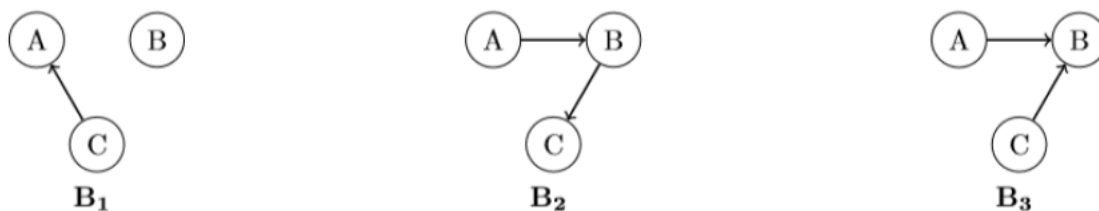
July 7th, 2021 at 11:59pm

### 1 Bayes' Net: Representation

Assume we are given the following ten Bayes' nets, labeled  $G_1$  to  $G_{10}$ :



Assume we are also given the following Bayes' nets, labeled  $B_1$  to  $B_3$ :



1. Assume we know that a joint distribution  $d_1$  (over A, B, C) can be represented by Bayes' net  $B_1$ .

Mark all of the following Bayes' nets that are guaranteed to be able to represent  $d_1$ .

- ☐  $G_1$     ☐  $G_2$     ☐  $G_3$     ☒  $G_4$     ☒  $G_5$   
☐  $G_6$     ☒  $G_7$     ☐  $G_8$     ☒  $G_9$     ☒  $G_{10}$

☐ None of the above.

2. Assume we know that a joint distribution  $d_2$ (over A,B,C) can be represented by Bayes' net **B<sub>2</sub>**. Mark all of the following Bayes' nets that are guaranteed to be able to represent  $d_2$ .

☐ **G<sub>1</sub>**      ☐ **G<sub>2</sub>**      ☐ **G<sub>3</sub>**      ☐ **G<sub>4</sub>**      ☐ **G<sub>5</sub>**  
☒ **G<sub>6</sub>**      ☐ **G<sub>7</sub>**      ☒ **G<sub>8</sub>**      ☒ **G<sub>9</sub>**      ☒ **G<sub>10</sub>**

☐ None of the above.

3. Assume we know that a joint distribution  $d_3$ (over A,B,C) can be represented by Bayes' net **B<sub>3</sub>**. Mark all of the following Bayes' nets that are guaranteed to be able to represent  $d_3$ .

☐ **G<sub>1</sub>**      ☐ **G<sub>2</sub>**      ☐ **G<sub>3</sub>**      ☐ **G<sub>4</sub>**      ☐ **G<sub>5</sub>**  
☐ **G<sub>6</sub>**      ☐ **G<sub>7</sub>**      ☐ **G<sub>8</sub>**      ☒ **G<sub>9</sub>**      ☒ **G<sub>10</sub>**

☐ None of the above.

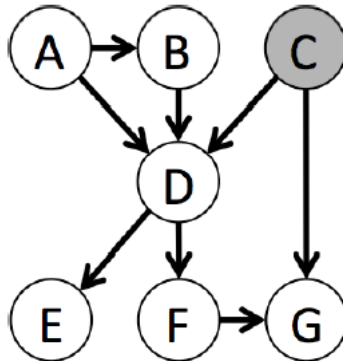
4. Assume we know that a joint distribution  $d_4$ (over A,B,C) can be represented by Bayes' net **B<sub>1</sub>B<sub>2</sub>** and **B<sub>3</sub>**. Mark all of the following Bayes' nets that are guaranteed to be able to represent  $d_4$ .

☒ **G<sub>1</sub>**      ☒ **G<sub>2</sub>**      ☒ **G<sub>3</sub>**      ☒ **G<sub>4</sub>**      ☒ **G<sub>5</sub>**  
☒ **G<sub>6</sub>**      ☒ **G<sub>7</sub>**      ☒ **G<sub>8</sub>**      ☒ **G<sub>9</sub>**      ☒ **G<sub>10</sub>**

☐ None of the above.

## 2 Variable Elimination

For the Bayes' net below, we are given the query  $P(A, E | +c)$ . All variables have binary domains. Assume we run variable elimination to compute the answer to this query, with the following variable elimination ordering: B, D, G, F.



Complete the following description of the factors generated in this process:

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

**Solution:**

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

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

**Solution:**

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

This leaves us with the factors:

**Solution:**

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

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

**Solution: ...**  $f_2(A, +c, E, F) = \sum_d f_1(A, +c, d) P(E | d) P(F | d)$

This leaves us with the factors:

$$\text{Solution: ... } P(A), P(+C), P(G|+C, F), f_2(A, +C, E, F)$$

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

$$\text{Solution: ... } f_3(+C, F) = \sum_g P(g|+C, F)$$

This leaves us with the factors:

$$\text{Solution: ... } P(A), P(+C), f_2(A, +C, E, F), f_3(+C, F)$$

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

$$\text{Solution: ... } f_4(A, +C, E) = \sum_f f_2(A, +C, E, f) f_3(+C, f)$$

This leaves us with the factors:

$$\text{Solution: ... } P(A), P(+C), f_4(A, +C, E)$$

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

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

(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.

$$\text{Solution: ... } f_2(A, +C, E, F), \text{ the size is } 2^3 = 8$$

(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,  $f_1(A, +c, D)$ .