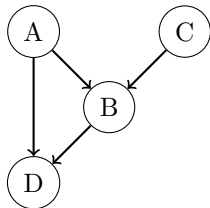


## Q1. Bayes Nets: Representation



Consider the Bayes net graph depicted above.

- (a) (i) Select all conditional independences that are enforced by this Bayes net graph.

- ☐  $A \perp\!\!\!\perp B$ 
☐  $A \perp\!\!\!\perp C \mid B$   
☐  $D \perp\!\!\!\perp C \mid A, B$ 
☐  $D \perp\!\!\!\perp C$   
☐  $A \perp\!\!\!\perp C$ 
☐  $A \perp\!\!\!\perp C \mid D$   
☐  $A \perp\!\!\!\perp C \mid B, D$ 
☐  $D \perp\!\!\!\perp C \mid B$

- (ii) Because of these conditional independences, there are some distributions that cannot be represented by this Bayes net. What is the minimum set of edges that would need to be added such that the resulting Bayes net could represent any distribution?

- ☐  $A \rightarrow C$ 
☐  $C \rightarrow A$   
☐  $C \rightarrow D$ 
☐  $D \rightarrow C$   
☐  $D \rightarrow A$ 
☐  $D \rightarrow B$   
☐  $B \rightarrow C$ 
☐  $B \rightarrow A$

- (b) For the rest of this Q2, we use the **original, unmodified** Bayes net depicted at the beginning of the problem statement. Here are some partially-filled conditional probability tables on  $A$ ,  $B$ ,  $C$ , and  $D$ . Note that these are not necessarily factors of the Bayes net. Fill in the six blank entries such that this distribution can be represented by the Bayes net.

$A$	$C$	$P(C \mid A)$
$+a$	$+c$	0.8
$+a$	$-c$	0.2
$-a$	$+c$	0.8
$-a$	$-c$	0.2

$A$	$B$	$D$	$P(D \mid A, B)$
$+a$	$+b$	$+d$	0.60
$+a$	$+b$	$-d$	0.40
$+a$	$-b$	$+d$	0.10
$+a$	$-b$	$-d$	0.90
$-a$	$+b$	$+d$	0.20
$-a$	$+b$	$-d$	0.80
$-a$	$-b$	$+d$	0.50
$-a$	$-b$	$-d$	0.50

$A$	$B$	$C$	$P(C \mid A, B)$
$+a$	$+b$	$+c$	0.50
$+a$	$+b$	$-c$	0.50
$+a$	$-b$	$+c$	0.20
$+a$	$-b$	$-c$	0.80
$-a$	$+b$	$+c$	0.90
$-a$	$+b$	$-c$	0.10
$-a$	$-b$	$+c$	0.40
$-a$	$-b$	$-c$	0.60

$C$	$P(C)$
$+c$	(i)
$-c$	(ii)

$A$	$B$	$C$	$D$	$P(D, C \mid A, B)$
$+a$	$+b$	$+c$	$+d$	(iii)
$+a$	$+b$	$-c$	$-d$	(iv)
$+a$	$-b$	$+c$	$+d$	(v)
$+a$	$-b$	$-c$	$-d$	(vi)
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

## Q2. Bayes Nets: Inference

Consider the following Bayes Net, where we have observed that  $D = +d$ .

$P(A)$		$P(B A)$			$P(C A, B)$				$P(D C)$		
$+a$	0.5	$+a$	$+b$	0.5	$+a$	$+b$	$+c$	0.8	$+c$	$+d$	0.4
$-a$	0.5	$+a$	$-b$	0.5	$+a$	$+b$	$-c$	0.2	$+c$	$-d$	0.6
		$-a$	$+b$	0.2	$+a$	$-b$	$+c$	0.6	$-c$	$+d$	0.2
		$-a$	$-b$	0.8	$-a$	$+b$	$+c$	0.2	$-c$	$-d$	0.8
					$-a$	$+b$	$-c$	0.8			
					$-a$	$-b$	$+c$	0.1			
					$-a$	$-b$	$-c$	0.9			

- (a) Below is a list of samples that were collected using prior sampling. Mark the samples that would be **rejected** by rejection sampling.

- ☐  $+a \quad -b \quad +c \quad -d$   
☐  $+a \quad -b \quad +c \quad +d$   
☐  $-a \quad +b \quad -c \quad -d$   
☐  $+a \quad +b \quad +c \quad +d$

- (b) To decouple from the previous part, you now receive a *new* set of samples shown below:

$+a \quad +b \quad +c \quad +d$   
 $-a \quad -b \quad -c \quad +d$   
 $+a \quad +b \quad +c \quad +d$   
 $+a \quad -b \quad -c \quad +d$   
 $-a \quad -b \quad -c \quad +d$

Estimate the probability  $P(+a|+d)$  if these new samples were collected using...

- (i) ... rejection sampling: \_\_\_\_\_  
 (ii) ... likelihood weighting: \_\_\_\_\_

- (c) Instead of sampling, we now wish to use **variable elimination** to calculate  $P(+a|+d)$ . We start with the factorized representation of the joint probability:

$$P(A, B, C, +d) = P(A)P(B|A)P(C|A, B)P(+d|C)$$

- (i) We begin by eliminating the variable  $B$ , which creates a new factor  $f_1$ . Complete the expression for the factor  $f_1$  in terms of other factors.

$$f_1(\text{_____}) = \text{_____}$$

- (ii) After eliminating  $B$  to create a factor  $f_1$ , we next eliminate  $C$  to create a factor  $f_2$ . What are the remaining factors after both  $B$  and  $C$  are eliminated?

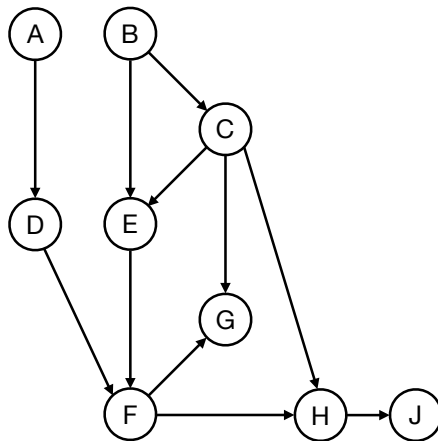
☐  $p(A)$       ☐  $p(B|A)$       ☐  $p(C|A, B)$       ☐  $p(+d|C)$       ☐  $f_1$       ☐  $f_2$

- (iii) After eliminating both  $B$  and  $C$ , we are now ready to calculate  $P(+a|+d)$ . Write an expression for  $P(+a|+d)$  in terms of the remaining factors.

$$P(+a|+d) = \text{_____}$$

### Q3. Bayes Nets: Elimination

(a) Consider running variable elimination on the Bayes Net shown below.



First, we eliminate  $D$  to create a factor  $f_1$   
 Next, we eliminate  $E$  to create a factor  $f_2$   
 Next, we eliminate  $H$  to create a factor  $f_3$

From the list below, select all factors that remain after  $D$ ,  $E$  and  $H$  have been eliminated.

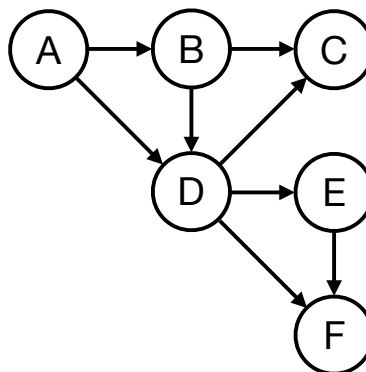
- ☐  $f_1$
- ☐  $f_2$
- ☐  $f_3$

- ☐  $P(A)$
- ☐  $P(A|F)$
- ☐  $P(B)$
- ☐  $P(B|A)$
- ☐  $P(B|C)$
- ☐  $P(C)$
- ☐  $P(C|B)$

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

- ☐  $P(G|C)$
- ☐  $P(G|C, F)$
- ☐  $P(H|F)$
- ☐  $P(H|C, F)$
- ☐  $P(H|J)$
- ☐  $P(J|C)$
- ☐  $P(J|F)$
- ☐  $P(J|H)$

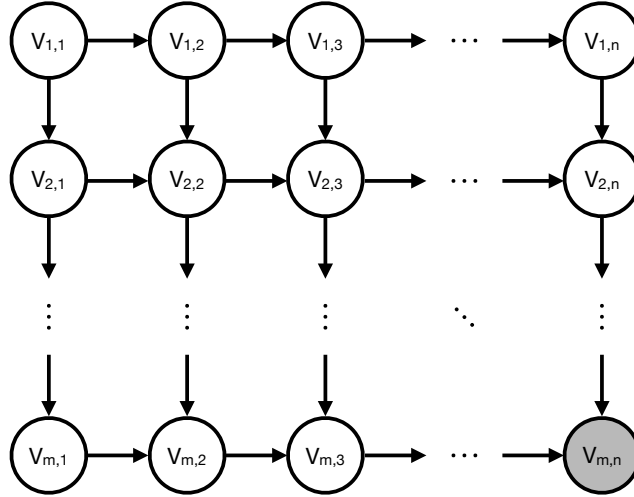
(b) Consider the Bayes Net shown below. Each variable in the Bayes Net can take on two possible values.



You are given the query  $P(C|F)$ , which you would like to answer using variable elimination. Please find a variable elimination ordering where the largest intermediate factor created during variable elimination is as small as possible.

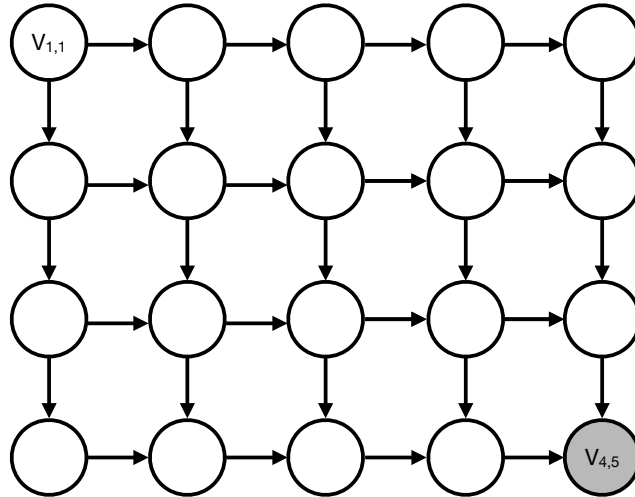
Elimination ordering: \_\_\_\_\_

- (c) Consider doing inference in an  $m \times n$  lattice Bayes Net, as shown below. The network consists of  $mn$  binary variables  $V_{i,j}$ , and you have observed that  $V_{m,n} = +v_{m,n}$ .



You wish to calculate  $P(V_{1,1} | v_{m,n})$  using variable elimination. To maximize computational efficiency, you wish to use a variable elimination ordering for which the size of the largest generated factor is as small as possible.

- (i) First consider the special case where  $m = 4$  and  $n = 5$ . A reproduction of the lattice is shown below, with variable names for non-query variables omitted. Please provide your optimal elimination ordering for this example by numbering the nodes below in the order they will be eliminated (i.e. write a number such as 1, 2, 3, ... inside every node that will be eliminated.)



- (ii) Now consider the general case (assume  $m > 2$  and  $n > 2$ ). What is the size of the largest factor generated under the most efficient elimination ordering? Your answer should be the number of rows in the factor's table, expressed in terms of  $m$  and  $n$ .

Size (number of rows) of the largest factor: \_\_\_\_\_