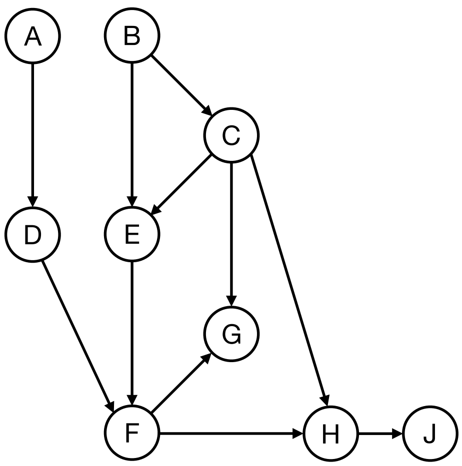
Introduction to Artificial Intelligence

Exam Prep 7 Solutions

Q1. Bayes Nets: Elimination

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

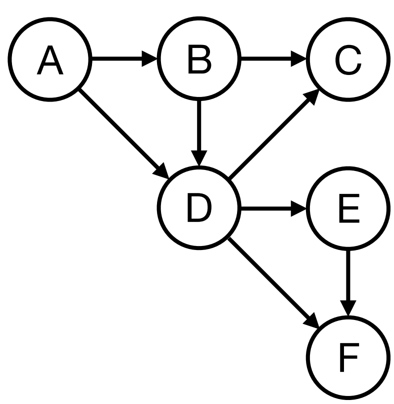


First, we eliminate D to create a factor f_1
Next, we eliminate E to create a factor f_2
Next, we eliminate H to create a factor f_3

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

- [ ] f_1
- [x] f_2
- [x] f_3
- [x] $P(A)$
- [ ] $P(A|F)$
- [x] $P(B)$
- [ ] $P(B|A)$
- [ ] $P(B|C)$
- [ ] $P(C)$
- [x] $P(C|B)$
- [ ] $P(D|A)$
- [ ] $P(D|A,F)$
- [ ] $P(E|B)$
- [ ] $P(E|C)$
- [ ] $P(E|B,C)$
- [ ] $P(F)$
- [ ] $P(F|A,B)$
- [ ] $P(F|G)$
- [ ] $P(G|C)$
- [x] $P(G|C,F)$
- [ ] $P(H|F)$
- [ ] $P(H|C,F)$
- [ ] $P(H|J)$
- [ ] $P(J|C)$
- [ ] $P(J|F)$
- [ ] $P(J|H)$

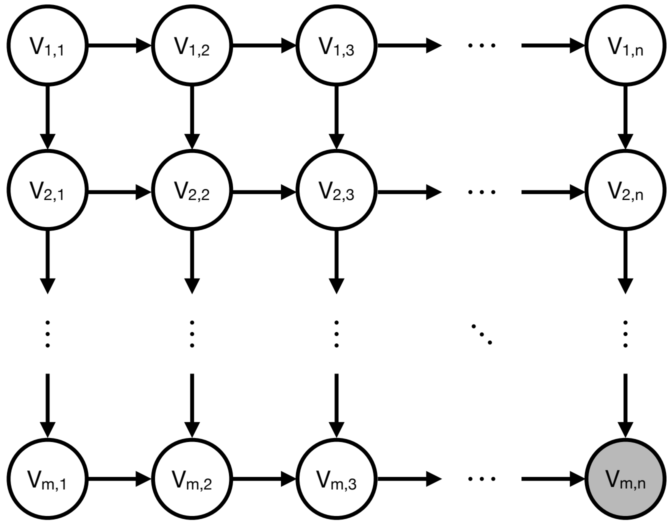
(b) Consider the Bayes Net shown below. Each variable in the Bayes Net can take on two possible values.



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.

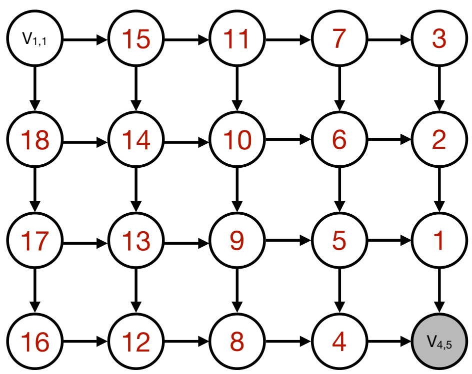
Elimination ordering: ABED, AEBD, or EABD

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

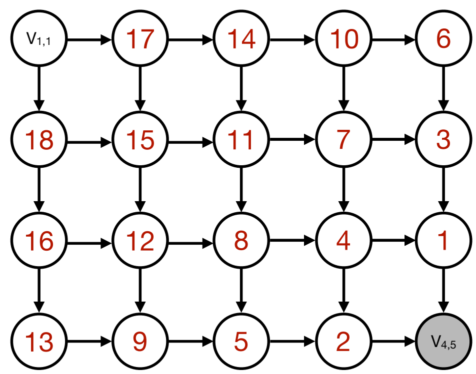


You wish to calculate $P(V_{1,1} \mid +v_{m,n})$ using variable elimination. To maximize computational efficiency, you wish to use a variable elimination ordering for which the size of the largest generated factor is as small as possible.

(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 $1, 2, 3, \ldots$ inside every node that will be eliminated.)

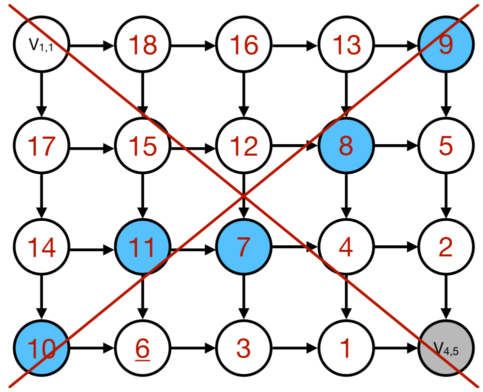


(or)



Note that there is actually more than one correct ordering, and that a few minor variations on the orderings given above are possible. However, it's important to start near the same corner as the evidence variable and to never create a factor that involves more than 4 non-evidence variables.

However, the ordering shown below is suboptimal (eliminating node 6 will create a size 2^5 factor involving the five nodes highlighted in blue):

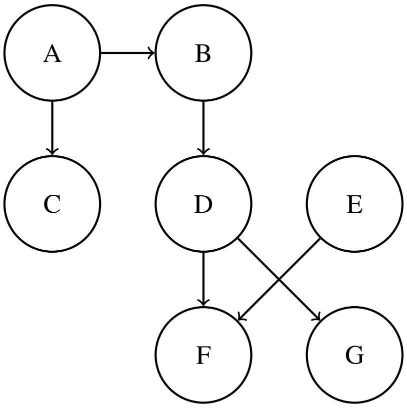


(ii) Now consider the general case (assume $m > 2$ and $n > 2$). What is the size of the largest factor generated under the most efficient elimination ordering? Your answer should be the number of rows in the factor's table, expressed in terms of m and n.

Size (number of rows) of the largest factor: $2^{\min(m,n)}$

Q2. Bayes' Nets: Representation and Independence

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



(a) Mark the statements that are guaranteed to be true.

- [ ] $B \perp\!\!\!\perp C$
- [ ] $A \perp\!\!\!\perp F$
- [ ] $D \perp\!\!\!\perp E | F$
- [x] $E \perp\!\!\!\perp A | D$
- [x] $F \perp\!\!\!\perp G | D$
- [x] $B \perp\!\!\!\perp F | D$
- [ ] $C \perp\!\!\!\perp G$
- [x] $D \perp\!\!\!\perp E$

Parts (b) and (c) pertain to the following probability distribution tables. The joint distribution $P(A,B,C,D)$ is equal to the product of these probability distribution tables.

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

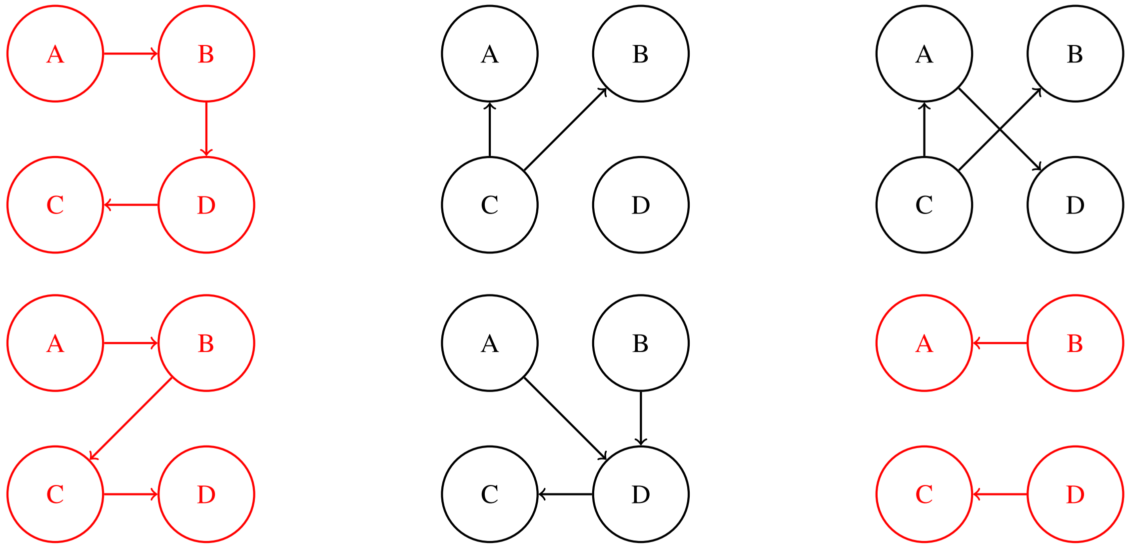
B	C	$P(C\|B)$
+b	+c	0.8
+b	–c	0.2
–b	+c	0.8
–b	–c	0.2

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

(b) 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 $C \not\perp\!\!\!\perp D$. Then, we have every remaining pair of variables: $A \perp\!\!\!\perp C, A \perp\!\!\!\perp D, B \perp\!\!\!\perp C, B \perp\!\!\!\perp D$

(c) 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 cause), so we cannot circle it.

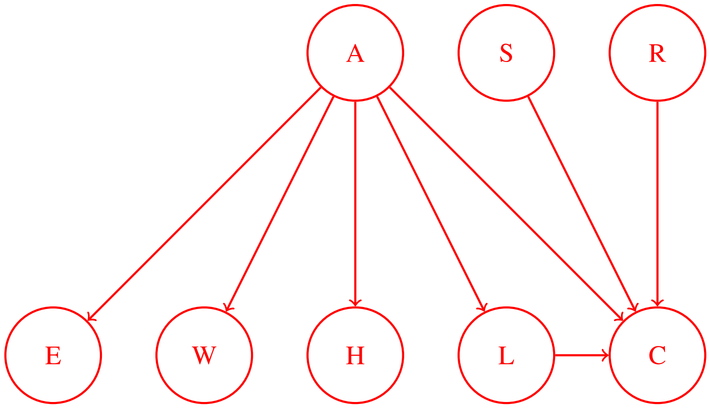
The top right net seems like it could represent the distribution, because D-separation finds that: A and B are not guaranteed to be independent (common effect), and C and D are not guaranteed to be independent (causal chain). However, according to Part D, $A \perp\!\!\!\perp C$, $A \perp\!\!\!\perp D$, and $B \perp\!\!\!\perp C$, so all of the arrows in the net are vacuous. That means, in this net, A and B are independent, and C and D are independent, so we cannot circle this net.

You are building advanced safety features for cars that can warn a driver if they are falling asleep (A) and also calculate the probability of a crash (C) in real time. You have at your disposal 6 sensors (random variables):

- E: whether the driver's eyes are open or closed
- W: whether the steering wheel is being touched or not
- L: whether the car is in the lane or not
- S: whether the car is speeding or not
- H: whether the driver's heart rate is somewhat elevated or resting
- R: whether the car radar detects a close object or not

A influences $\{E, W, H, L, C\}$. C is influenced by $\{A, S, L, R\}$.

(d) Draw the Bayes Net associated with the description above by adding edges between the provided nodes where appropriate.



(e) Mark all the independence assumptions that must be true.

- [x] $E \perp\!\!\!\perp S$
- [x] $W \perp\!\!\!\perp H|A$
- [x] $S \perp\!\!\!\perp R$
- [ ] $E \perp\!\!\!\perp L$
- [ ] $L \perp\!\!\!\perp R|C$
- [x] $W \perp\!\!\!\perp R$
- [ ] $A \perp\!\!\!\perp C$
- [ ] $E \perp\!\!\!\perp C|L$

(f) The car's sensors tell you that the car is in the lane ($L = +l$) and that the car is not speeding ($S = -s$). Now you would like to calculate the probability of crashing, $P(C|+l, -s)$. We will use the variable elimination ordering R, A, E, W, H. Write down the largest factor generated during variable elimination. Box your answer.

Our factors if we don't observe evidence are $P(A)$, $P(S)$, $P(R)$, $P(E|A)$, $P(W|A)$, $P(H|A)$, $P(L|A)$, $P(C|L, A, S, R)$. We observe evidence, and we have: $P(A)$, $P(R)$, $P(E|A)$, $P(W|A)$, $P(H|A)$, $P(C|+l, A, -s, R)$. We first eliminate R, so we select $P(R)$ and $P(C|+l, A, -s, R)$ to get $f_1(C|+l, A, -s)$. Now we eliminate A, so we select $P(A)$, $P(E|A)$, $P(W|A)$, $P(H|A)$, $P(+l|A)$, $f_1(C|+l, A, -s)$ and get $\boxed{f_2(C, E, W, H|+l, -s)}$. We see that this must be the largest factor because this is the only factor we have left at this point, and variable elimination is not yet finished.

(g) Write down a more efficient variable elimination ordering, i.e. one whose largest factor is smaller than the one generated in the previous question.

Any ordering of the five variables where at least one of $\{E, W, H\}$ is before A would be more efficient than the previous ordering. As an example, R, E, W, H, A would work.