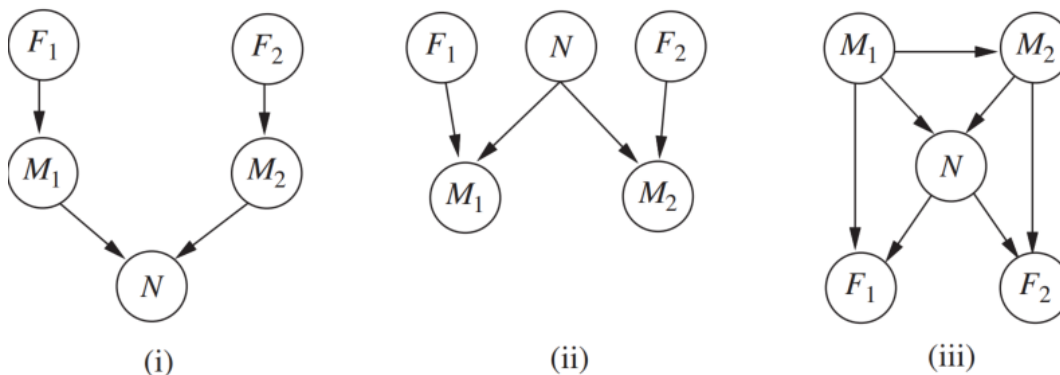


# CSCI 390, Artificial Intelligence

## Assignment 4

### Question 1 (Bayes Net representation)

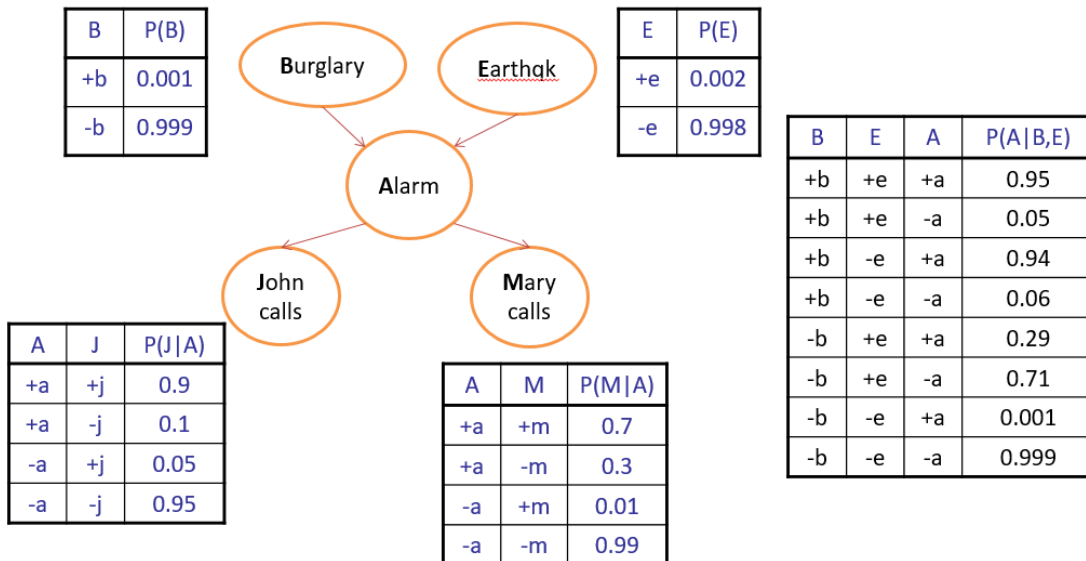
Two astronomers in different parts of the world make measurements  $M_1$  and  $M_2$  of the number of stars  $N$  in some small region of the sky, using their telescopes. Normally, there is a small possibility  $e$  of error by up to one star in each direction. Each telescope can also (with a much smaller probability  $f$ ) be badly out of focus (events  $F_1$  and  $F_2$ ), in which case the scientist will undercount by three or more stars (or if  $N$  is less than 3, fail to detect any stars at all). Consider the three networks shown in the following figure.



- (a) Which of these Bayesian networks are correct (but not necessarily efficient) representations of the preceding information?
- (b) Which is the best network? Explain.
- (c) Write out a conditional distribution for  $\mathbf{P}(M_1 | N)$ , for the case where  $N \in \{1, 2, 3\}$  and  $M_1 \in \{0, 1, 2, 3, 4\}$ . Each entry in the conditional distribution should be expressed as a function of the parameters  $e$  and/or  $f$ .
- (d) Suppose  $M_1 = 1$  and  $M_2 = 3$ . What are the *possible* numbers of stars if you assume no prior constraint on the values of  $N$ ?
- (e) What is the *most likely* number of stars, given these observations? Explain how to compute this, or if it is not possible to compute, explain what additional information is needed and how it would affect the result.

## Question 2 (Exact inference)

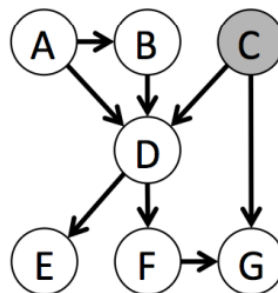
Consider the alarm network shown in the lecture as following:



- Compute  $P(B|+j, +m)$  by using variable elimination. Note: perform all the necessary steps in variable elimination.
- Count the number of arithmetic operations performed, and compare it with the number performed by the enumeration algorithm.

## Question 3 (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:

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

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

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

This leaves us with the factors:

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

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

This leaves us with the factors:

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

This leaves us with the factors:

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

This leaves us with the factors:

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

(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.

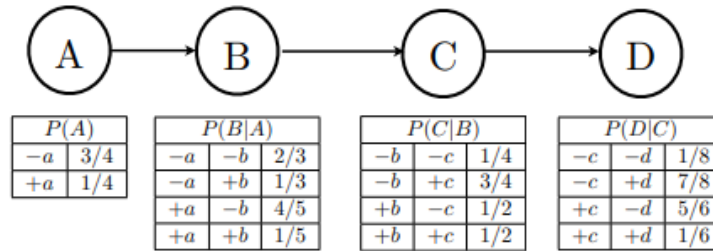
(d) Find a variable elimination ordering for the same query, i.e., for  $P(A, E | +c)$ , for which the maximum size factor generated along the way is smallest. Hint: the maximum size factor generated in your solution should have only 2 variables, for a size of  $2^2 = 4$  table. Fill in the variable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

Variable Eliminated	Factor Generated

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

### Question 4 (Sampling)

Assume the following Bayes' net, and the corresponding distributions over the variables in the Bayes' net:



- (a) You are given the following samples:
- |                  |                  |
|------------------|------------------|
| $+a + b - c - d$ | $+a - b - c + d$ |
| $+a - b + c - d$ | $+a + b + c - d$ |
| $-a + b + c - d$ | $-a + b - c + d$ |
| $-a - b + c - d$ | $-a - b + c - d$ |
- Assume that these samples came from performing **Prior Sampling**, and calculate the sample estimate of  $P(+c)$ .
  - Now we will estimate  $P(+c \mid +a, -d)$ . Above, clearly cross out the samples that would not be used when doing **Rejection Sampling** for this task, and write down the sample estimate of  $P(+c \mid +a, -d)$ .
- (b) Using **Likelihood Weighting Sampling** to estimate  $P(-a \mid +b, -d)$ , the following samples were obtained (shown in the column of Sample). Fill in the weight of each sample in the corresponding row of the below table.

Sample	Weight
$-a + b + c - d$	
$+a + b + c - d$	
$+a + b - c - d$	
$-a + b - c - d$	

- (c) From the weighted samples in the previous question, estimate  $P(-a \mid +b, -d)$ .
- (d) Which query is better suited for likelihood weighting,  $P(D \mid A)$  or  $P(A \mid D)$ ? Justify your answer in one sentence.