

Chapter 14 (AIAMA)

Probabilistic Reasoning

Sukarna Barua

Assistant Professor, CSE, BUET

Bayesian Network: Exact Inference

- **Consider the query:** $P(\text{Burglary} | \text{JohnCalls} = \text{ture}, \text{MaryCalls} = \text{true})$
- We need to compute: $P(B | j, m)$ [compute for both b and $\neg b$]
- Express in terms of joint distributions [we already know how to compute joint distribution]

$$P(B | j, m) = \frac{P(B, j, m)}{P(j, m)} = \alpha P(B, j, m)$$

where α is the normalizing constant [we can find the value of α later]

Bayesian Network: Exact Inference

- How to compute $P(B, j, m)$?
- Add all hidden variables to get full joint distribution

$$P(B, j, m) = \sum_a \sum_e P(B, j, m, a, e) = \sum_a \sum_e P(j|a)P(m|a)P(a|B, e)P(B)P(e)$$

- Query variable: B
- Evidence variables: J, M
- Hidden variables: A, E

Bayesian Network: Exact Inference

- Finally, we get the following:

$$P(B|j, m) = \alpha \sum_a \sum_e P(j|a)P(m|a)P(a|B, e)P(B)P(e)$$

- Compute $P(b|j, m)$ and $P(\neg b|j, m)$ [expression has α]
 - Find α [Using the eq. $P(b|j, m) + P(\neg b|j, m) = 1$]
- **Question:** How many operations (multiplications + additions)?
 - Answer: [For each $P(B|j, m)$] Mult: 16, Sum: 4, Total = 20 operations

Bayesian Network: Exact Inference

- Computing using joint distribution naively:

$$P(B|j, m) = \alpha \sum_a \sum_e P(j|a)P(m|a)P(a|B, e)P(B)P(e)$$

- Required operations: 20
- **Can we improve?**
 - Yes, move sums inside [*closest to the factors having the hidden variable*]
 - Sums are evaluated early and gets multiplied once [*rather than every time during joint calculation*]

Bayesian Network: Exact Inference

- Move summations inside: $P(B|j,m) = \alpha \sum_a \sum_e P(j|a)P(m|a)P(a|B,e)P(B)P(e)$
- Two alternate ways to do it:
 - Option 1: $P(B|j,m) = \alpha P(B) \sum_a P(j|a)P(m|a) \sum_e P(a|B,e)P(e)$
 - Option 2: $P(B|j,m) = \alpha P(B) \sum_e P(e) \sum_a P(a|B,e)P(j|a)P(m|a)$
- Which one is more efficient in terms of operations?

Bayesian Network: Exact Inference

- **Which one is more efficient?** Assume $B = \text{true}$ and compute both.
- Equation 1: $P(B|j,m) = \alpha P(B) \sum_a P(j|a)P(m|a) \sum_e P(a|B,e)P(e)$
 - Required # of multiplications: 9, # of sums = 3, total = 12
- Equation 2: $P(B|j,m) = \alpha P(B) \sum_e P(e) \sum_a P(a|B,e)P(j|a)P(m|a)$
 - Required # of multiplications: 11, # of sums = 3, total = 14
- So order matters! However, finding optimal order is difficult!

Bayesian Network: Exact Inference

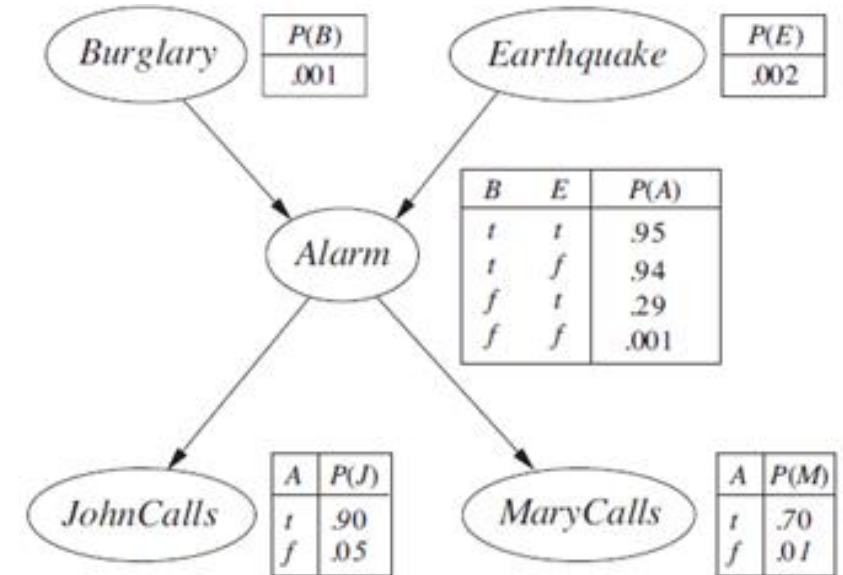
- Using the CPT data, we calculate:

$$P(b|j, m) = \alpha \times 0.00059224$$

$$P(\neg b|j, m) = \alpha \times 0.0014919$$

Hence,

$$\begin{aligned} P(B|j, m) &= \alpha \times (0.00059224, 0.0014919) \\ &= (0.284, 0.716) \text{ [After normalizing as probabilities sum to 1]} \end{aligned}$$

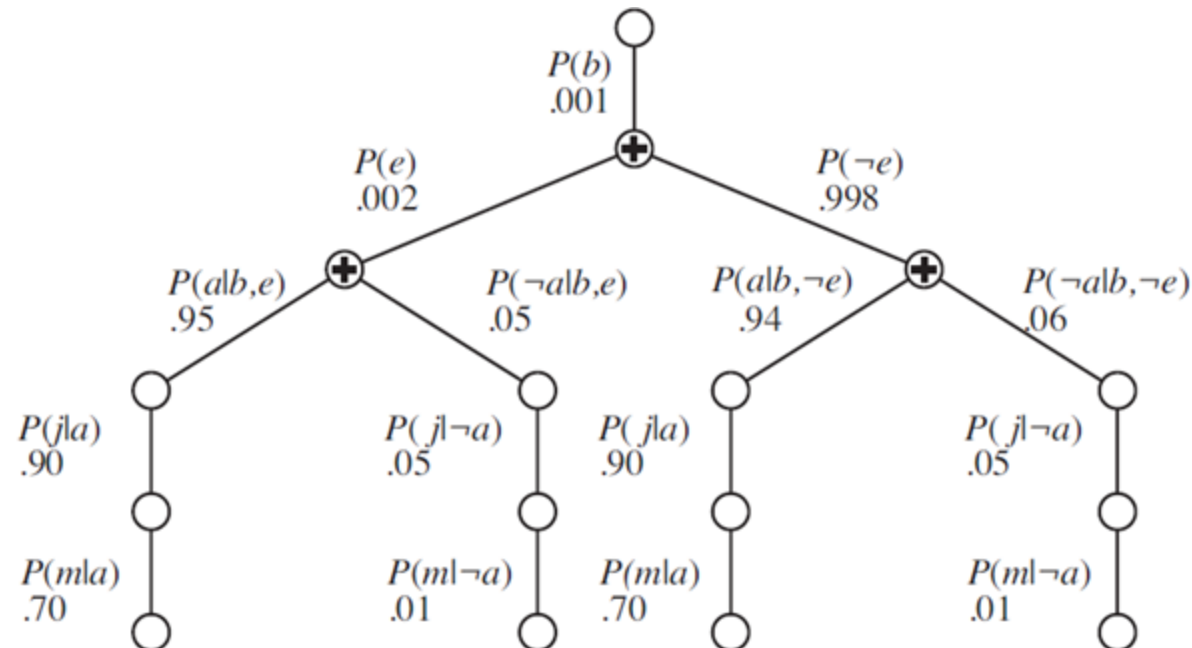


Bayesian Network: Exact Inference

- An evaluation tree is shown for the expression:

$$\alpha P(B) \sum_e P(e) \sum_a P(a|B, e) P(j|a) P(m|a)$$

- Blank circled nodes represent multiplication [Notice the repetition of sub-paths]

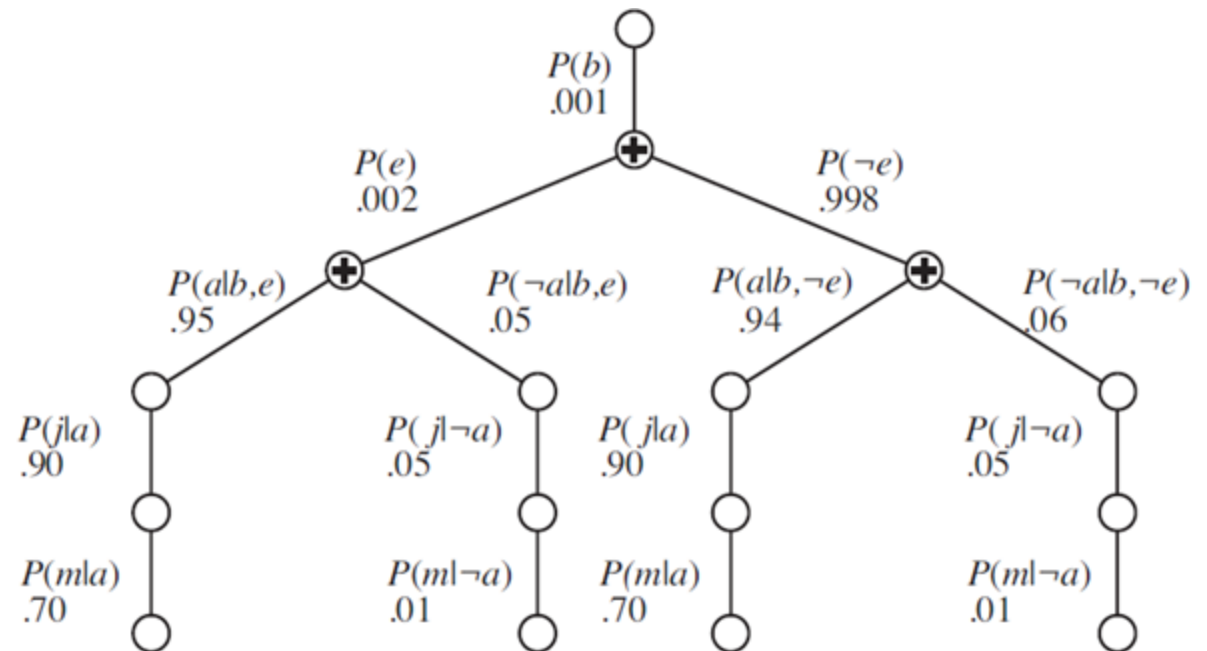


Bayesian Network: Exact Inference

- **Problem:** Repeated subexpression evaluation
- The following products are computed twice in two paths of the tree!

$$P(j \mid a)P(m \mid a)$$

$$P(j \mid \neg a)P(m \mid \neg a)$$



Bayesian Network: Variable Elimination Algorithm

- Steps of variable elimination algorithm:
 - Step 1: Compute factors for each node of Bayesian Network and instantiate/restrict evidence variables.
 - Step 2: For some order of variables (V_1, V_2, \dots, V_N) :
 - (2a) If variable V is a hidden variable, then multiply all the factors of V
 - (2b) Sum-out V
 - Step 3: Multiply remaining factors.
 - Step 4: Normalize the final factor to eliminate α .

Bayesian Network: Variable Elimination Algorithm

- Consider the following expression: $P(B|j, m)$.
 - Query variable: B
 - Evidence variable: J, M
 - Hidden variable: A, E
- We will use variable elimination algorithm to calculate

Bayesian Network: Variable Elimination Algorithm

- Query: $P(B|j, m)$.
- Joint probability expression of BN:

$$P(B, E, A, J, M) = P(B)P(E)P(A|B, E)P(J|A)P(M|A)$$

- **Step 1:** Compute all the factors of BN.
- The BN has five factors [*note the joint probability expression*]:
 - $f_1(B) = P(B)$
 - $f_2(E) = P(E)$
 - $f_3(A, B, E) = P(A|B, E)$
 - $f_4(J, A) = P(J|A)$
 - $f_5(M, A) = P(M|A)$

Variable Elimination Algorithm: Compute Factors

- Query: $P(B|j, m)$.
- **Step 1a:** Compute all the factors of BN. [*don't restrict evidences now*]

$$f_1(B) = P(B)$$

T	0.001
F	0.999

$$f_2(E) = P(E)$$

T	0.002
F	0.998

$$f_3(A, B, E) = P(A|B, E)$$

T	T	T	0.95
T	T	F	0.94
T	F	T	0.29
T	F	F	0.001
F	T	T	0.05
F	T	F	0.06
F	F	T	0.71
F	F	F	0.999

$$f_4(J, A) = P(J|A)$$


T	T	0.90
T	F	0.05
F	T	0.10
F	F	0.95

$$f_5(M, A) = P(M|A)$$

T	T	0.70
T	F	0.01
F	T	0.30
F	F	0.99


Variable Elimination Algorithm: Restrict Factors

- Query: $P(B|j, m)$.
- **Step 1b:** Instantiate/restrict evidence variables in all factors [j and m].
 - Restrict J to true in $f_4(J, A)$. [*Keep only rows where $J = \text{true}$*]
 - Note that J is removed from factor expression.

$f_4(J, A) = P(J A)$				$f_4(A)$	
T	T	0.90		T	0.90
T	F	0.05		F	0.05
F	T	0.10			
F	F	0.95			

Variable Elimination Algorithm: Restrict Factors

- Query: $P(B|j, m)$.
- **Step 1b:** Instantiate/restrict evidence variables in all factors. [j and m].
 - Restrict M to true in $f_5(M, A)$. [*Keep only rows where $M = \text{true}$*]

$f_5(M, A) = P(M A)$					$f_5(A)$
T	T	0.70		T	0.70
T	F	0.01		F	0.01
F	T	0.30			
F	F	0.99			

Variable Elimination Algorithm: Multiply Factors

- Query: $P(B|j, m)$.
- **Step 2:** Select every variable V (in any order) which is hidden [Here, A and E]
 - Multiply all factors of V and then sum-out the variable.
 - Suppose we select in this order: E, A .
 - First we select E
 - Multiple all factors having E and get a new factor: $f_6(A, B, E) = f_2(E) \times f_3(A, B, E)$
 - Sum-out E and get a new factor: $f_7(A, B) = \sum_E f_6(A, B, E)$
 - Note that after summing-out, E is removed from factor-argument.

Bayesian Network: Multiply Factors

- **Step 2a:** Multiple all factors having E and get a new factor:

$$f_6(A, B, E) = f_2(E) \times f_3(A, B, E)$$

- Factor multiplication is a point-wise multiplication of elements in two factors:
 - Multiply rows which have matching values of the variables.
 - Number of rows in the new factor 2^N where N is number of total variables in the new factor:
 - Multiplying $f_1(A, B)$ and $f_2(B, C)$ will give us $f_3(A, B, C)$ which will have 8 rows.

Bayesian Network: Multiply Factors

- For variable E :
 - Step 2a:** Multiple all factors having E and get a new factor:
 - Multiply:** $f_6(A, B, E) = f_2(E) \times f_3(A, B, E)$

$$f_2(E) = P(E)$$

T	0.002
F	0.998

×

$$f_3(A, B, E) = P(A|B, E)$$

T	T	T	0.95
T	T	F	0.94
T	F	T	0.29
T	F	F	0.001
F	T	T	0.05
F	T	F	0.06
F	F	T	0.71
F	F	F	0.999



$$f_6(A, B, E)$$

T	T	T	0.0019
T	T	F	0.93812
T	F	T	0.00058
T	F	F	0.000998
F	T	T	0.0001
F	T	F	0.05988
F	F	T	0.00142
F	F	F	0.997002

Bayesian Network: Multiply Factors

- For variable E:
 - **Step 2b:** Sum-out resulting factor for E and get a new factor:
 - **Sum-out E:** $f_7(A, B) = \sum_E f_6(A, B, E)$ [*sum corresponding rows that have T and F for E*]

$$f_6(A, B, E)$$

T	T	T	0.0019
T	T	F	0.93812
T	F	T	0.00058
T	F	F	0.000998
F	T	T	0.0001
F	T	F	0.05988
F	F	T	0.00142
F	F	F	0.997002



$$f_7(A, B)$$

T	T	0.94002
T	F	0.001578
F	T	0.05998
F	F	0.998422

Bayesian Network: Multiply Factors

- For variable A:
 - Step 2a:** Multiple all factors having A and get a new factor:
 - Multiply:** $f_8(A, B) = f_4(A) \times f_5(A) \times f_7(A, B)$

$f_4(A)$

T	0.90
F	0.05

\times

$f_5(A)$

T	0.70
F	0.01

\Rightarrow

$f_4(A) \times f_5(A)$

T	0.63
F	0.0005

\times

$f_7(A, B)$

T	T	0.94002
T	F	0.001578
F	T	0.05998
F	F	0.998422

\Rightarrow

$f_8(A, B)$

T	T	0.592213
T	F	0.00099414
F	T	0.00002999
F	F	0.000499211

Bayesian Network: Multiply Factors

- For variable A:
 - **Step 2b:** Sum-out resulting factor for E and get a new factor:
 - **Sum-out A:** $f_9(B) = \sum_A f_8(A, B)$ [*sum corresponding rows that have T and F for E*]

$f_8(A, B)$					$f_9(B)$
T	T	0.592213	⇒	T	0.59224299
T	F	0.00099414		F	0.001493351
F	T	0.00002999			
F	F	0.000499211			

Bayesian Network: Multiply Remaining Factors

- Query: $P(B|j, m)$.
- **Step 3:** Multiply remaining factors: $f_{10}(B) = f_1(B) \times f_9(B)$

$$f_1(B) = P(B) \quad \times \quad f_9(B) \quad \Rightarrow \quad f_{10}(B)$$

T	0.001
F	0.999

T	0.59224299
F	0.001493351

T	0.000592
F	0.001492

Bayesian Network: Normalize Result

- Query: $P(B|j, m)$.
- **Step 4:** Normalize the final factor to eliminate α .

$f_{10}(B)$			$f_{11}(B)$	
T	0.000592	\Rightarrow	T	0.284172
F	0.001492		F	0.715828

- **Answer:** $P(b|m, j) = 0.284, P(\neg b|m, j) = 0.716$

Bayesian Network: Variable Elimination Algorithm

- Algorithm pseudocode [RN book]
 - Note: Step 1 (initialization) is missing in the pseudocode [it is implied in the MAKE-FACTOR function which is not a good approach!]

```
function ELIMINATION-ASK( $X, \mathbf{e}, bn$ ) returns a distribution over  $X$   
  inputs:  $X$ , the query variable  
            $\mathbf{e}$ , observed values for variables  $\mathbf{E}$   
            $bn$ , a Bayesian network specifying joint distribution  $\mathbf{P}(X_1, \dots, X_n)$   
  
   $factors \leftarrow []$   
  for each  $var$  in ORDER( $bn.VARS$ ) do  
     $factors \leftarrow [MAKE-FACTOR(var, \mathbf{e}) | factors]$   
    if  $var$  is a hidden variable then  $factors \leftarrow SUM-OUT(var, factors)$   
  return NORMALIZE(POINTWISE-PRODUCT( $factors$ ))
```

Bayesian Network: Irrelevant Variable

- Consider the query: $P(\text{JohnCalls} | \text{Burglary} = \text{true})$
- Incorporating all hidden variables of the Bayesian Network, we get:

$$P(J|b) = \alpha P(J, b) = \alpha \sum_e \sum_a \sum_m P(j, m, a, b, e) = \sum_e \sum_a \sum_m P(J|a)P(m|a)P(a|b, e)P(b)P(e)$$

Bayesian Network: Irrelevant Variable

- A simple rearrangement of the expression would yield us:

$$\mathbf{P}(J | b) = \alpha P(b) \sum_e P(e) \sum_a P(a | b, e) \mathbf{P}(J | a) \sum_m P(m | a) .$$

- Note that the inner term $\sum_m P(m|a)$ will evaluate to 1. So it is eliminated from the expression entirely [Irrelevant variable]
- *Variable elimination algorithm will automatically sum-out m from the expression!*
 - Why? [At some point, m will be considered as a hidden variable and the summation over m will eliminate it]

Bayesian Network: Irrelevant Variable

- In general, we can remove any leaf node that is not a query variable or an evidence variable.
- After removal of initial leaves, there may be some more leaf nodes, and these too may be irrelevant and removed [continue this process].
- *Every variable that is not an ancestor of a query variable or evidence variable is irrelevant to the query.*