Exact Inference in Bayesian Networks

Outline

I. Probabilistic query using a BN

II. Variable Elimination

III. Variable ordering and relevance

* Figures are either from the textbook site.

I. Probabilistic Query

X: query variable

$\boldsymbol{E} = \{E_1, \ldots. E_m\}$: evidence variables

$\boldsymbol{e} = \{e_1, \ldots. e_m\}$: an observed event

$\boldsymbol{Y} = \{Y_1, \ldots. Y_l\}$: hidden variables

I. Probabilistic Query

X: query variable

$\boldsymbol{E} = \{E_1, \ldots . E_m\}$: evidence variables

$\boldsymbol{e} = \{e_1, \ldots . e_m\}$: an observed event

$\boldsymbol{Y} = \{Y_1, \ldots . Y_l\}$: hidden variables

Complete variable set: $\{X\} \cup \boldsymbol{E} \cup \boldsymbol{Y}$

I. Probabilistic Query

X: query variable

$\boldsymbol{E} = \{E_1, \ldots . E_m\}$: evidence variables

$\boldsymbol{e} = \{e_1, \ldots . e_m\}$: an observed event

$\boldsymbol{Y} = \{Y_1, \ldots . Y_l\}$: hidden variables

Complete variable set: $\{X\} \cup \boldsymbol{E} \cup \boldsymbol{Y}$

Query $\boldsymbol{P}(X \mid \boldsymbol{e})$?

I. Probabilistic Query

X: query variable

$\boldsymbol{E} = \{E_1, \ldots. E_m\}$: evidence variables

$\boldsymbol{e} = \{e_1, \ldots. e_m\}$: an observed event

$\boldsymbol{Y} = \{Y_1, \ldots. Y_l\}$: hidden variables

Complete variable set: $\{X\} \cup \boldsymbol{E} \cup \boldsymbol{Y}$

Query $\boldsymbol{P}(X \mid \boldsymbol{e})$?

// $\boldsymbol{P}(Burglary \mid j, m)$

$\boldsymbol{P}(Burglary \mid JohnCalls = true, MaryCalls = true) = \langle 0.284, 0.716 \rangle$

I. Probabilistic Query

X: query variable

$\boldsymbol{E} = \{E_1, \ldots. E_m\}$: evidence variables

$\boldsymbol{e} = \{e_1, \ldots. e_m\}$: an observed event

$\boldsymbol{Y} = \{Y_1, \ldots. Y_l\}$: hidden variables

Complete variable set: $\{X\} \cup \boldsymbol{E} \cup \boldsymbol{Y}$

Query $\boldsymbol{P}(X \mid \boldsymbol{e})$?

// $\boldsymbol{P}(\mathit{Burglary} \mid j, m)$

$\boldsymbol{P}(\mathit{Burglary} \mid \mathit{JohnCalls} = \mathit{true}, \mathit{MaryCalls} = \mathit{true}) = \langle 0.284, 0.716 \rangle$

We will discuss exact algorithms for posterior probability computation.

Inference by Enumeration

$$\mathbf{P}(X \mid \mathbf{e}) = \alpha \mathbf{P}(X, \mathbf{e}) = \alpha \sum_{\mathbf{y}} \boxed{\mathbf{P}(X, \mathbf{e}, \mathbf{y})}$$

Inference by Enumeration

$$\mathbf{P}(X \mid \mathbf{e}) = \alpha \mathbf{P}(X, \mathbf{e}) = \alpha \sum_{\mathbf{y}} \boxed{\mathbf{P}(X, \mathbf{e}, \mathbf{y})}$$

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid \mathit{Parents}(X_i))$$

Inference by Enumeration

$$\boldsymbol{P}(X \mid \boldsymbol{e}) = \alpha \boldsymbol{P}(X, \boldsymbol{e}) = \alpha \sum_{\boldsymbol{y}} \boxed{\boldsymbol{P}(X, \boldsymbol{e}, \boldsymbol{y})}$$

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid \mathit{Parents}(X_i))$$

computable as products of
conditional probabilities from the BN.

Inference by Enumeration

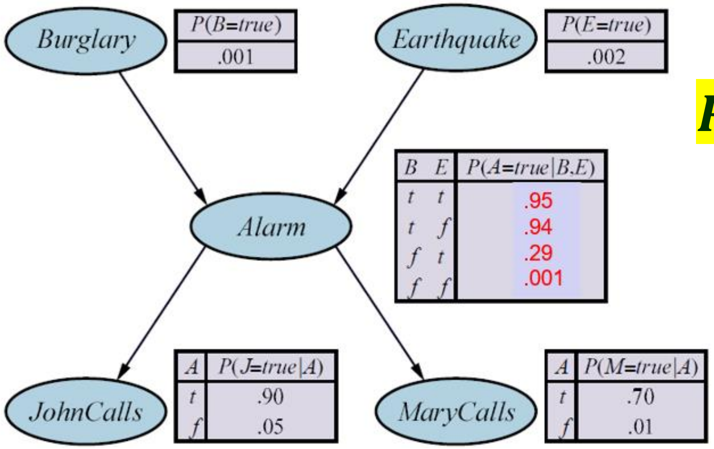
$$\boldsymbol{P}(X \mid \boldsymbol{e}) = \alpha \boldsymbol{P}(X, \boldsymbol{e}) = \alpha \sum_{\boldsymbol{y}} \boxed{\boldsymbol{P}(X, \boldsymbol{e}, \boldsymbol{y})}$$

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i))$$

computable as products of
conditional probabilities from the BN.

- Answer the query $\boldsymbol{P}(X \mid \boldsymbol{e})$ using a BN.

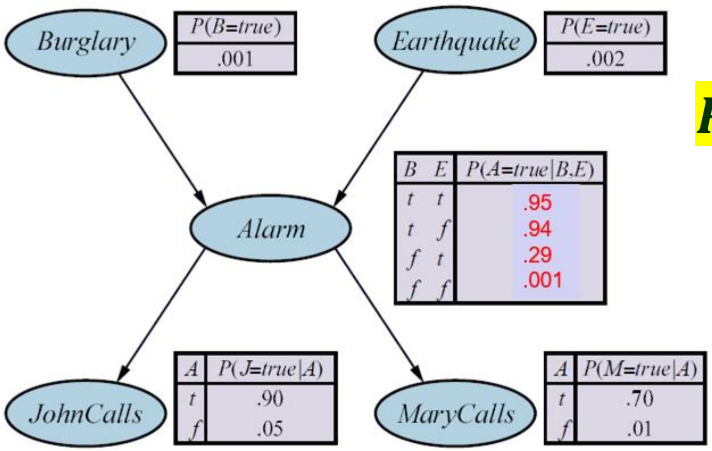
Burglary Example (revisited)



$\boldsymbol{P}(Burglary \mid JohnCalls = true, MaryCalls = true)?$

Hidden variables: $Y = \{Earthquake, Alarm\}$

Burglary Example (revisited)

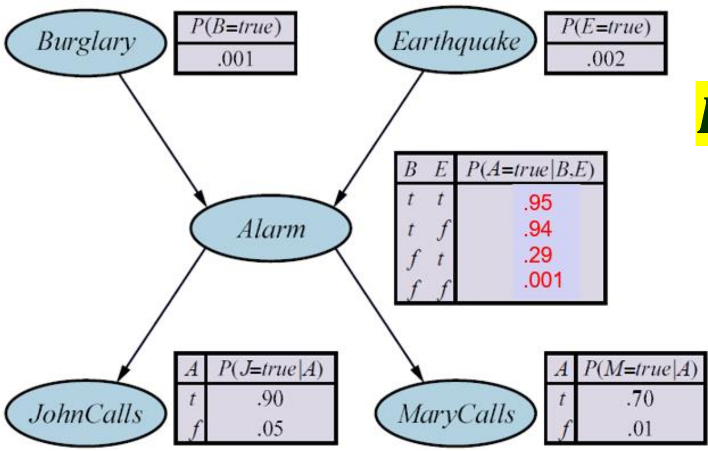


$\boldsymbol{P}(Burglary \mid JohnCalls = true, MaryCalls = true)$?

Hidden variables: $Y = \{Earthquake, Alarm\}$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{P}(B, j, m) = \alpha \sum_{E \in \{e, \neg e\}} \sum_{A \in \{a, \neg a\}} \boldsymbol{P}(B, j, m, E, A)$$

Burglary Example (revisited)



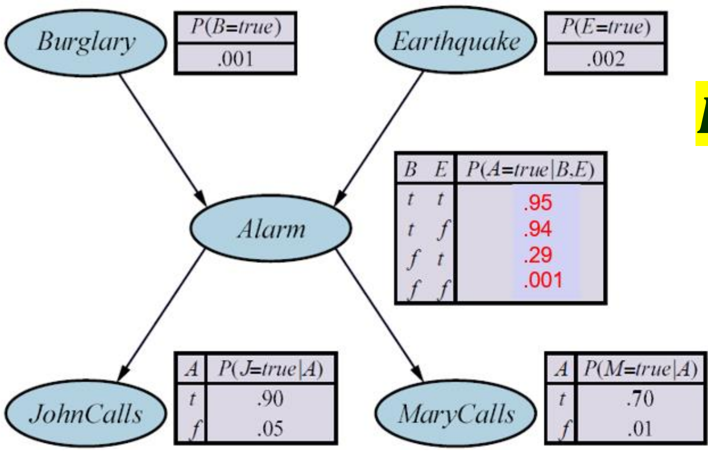
$\boldsymbol{P}(Burglary \mid JohnCalls = true, MaryCalls = true)$?

Hidden variables: $Y = \{Earthquake, Alarm\}$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{P}(B, j, m) = \alpha \sum_{E \in \{e, \neg e\}} \sum_{A \in \{a, \neg a\}} \boldsymbol{P}(B, j, m, E, A)$$

// the textbook uses e, a instead of E, A in the summation.
// this is incorrect because e is a constant value standing
// for *Earthquake* = *true*. the random variable E can take
// on both e and $\neg e$ needed for the summation. you may
// also call it an abuse of the notation e and a.

Burglary Example (revisited)



$\boldsymbol{P}(Burglary \mid JohnCalls = true, MaryCalls = true)$?

Hidden variables: $Y = \{Earthquake, Alarm\}$

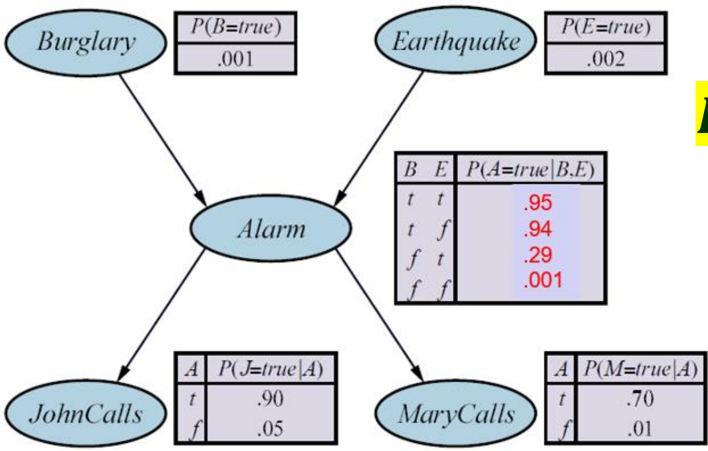
$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{P}(B, j, m) = \alpha \underbrace{\sum_{E \in \{e, \neg e\}} \sum_{A \in \{a, \neg a\}}} \boldsymbol{P}(B, j, m, E, A)$$

// the textbook uses e, a instead of E, A in the summation.
// this is incorrect because e is a constant value standing
// for $Earthquake = true$. the random variable E can take
// on both e and $\neg e$ needed for the summation. you may
// also call it an abuse of the notation e and a.

evaluated for b ($Burglary = true$) based on

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i))$$

Burglary Example (revisited)



$\boldsymbol{P}(\mathit{Burglary} \mid \mathit{JohnCalls} = \mathit{true}, \mathit{MaryCalls} = \mathit{true})$?

Hidden variables: $Y = \{\mathit{Earthquake}, \mathit{Alarm}\}$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{P}(B, j, m) = \alpha \underbrace{\sum_{E \in \{e, \neg e\}} \sum_{A \in \{a, \neg a\}}} \boldsymbol{P}(B, j, m, E, A)$$

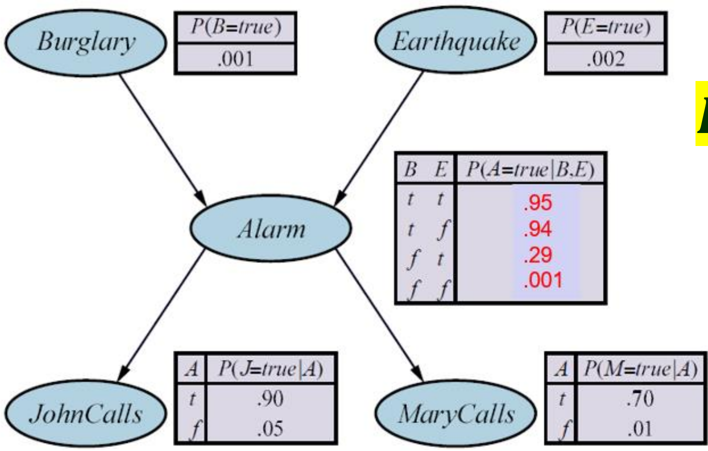
// the textbook uses e, a instead of E, A in the summation.
// this is incorrect because e is a constant value standing
// for *Earthquake = true*. the random variable E can take
// on both e and $\neg e$ needed for the summation. you may
// also call it an abuse of the notation e and a.

evaluated for b (*Burglary = true*) based on

$$P(x_1, \dots, x_n) = \prod_{i=1}^{n} P(x_i \mid \mathit{Parents}(X_i))$$

$$P(b \mid j, m) = \alpha \sum_{E} \sum_{A} P(b) P(E) P(A \mid b, E) P(j \mid A) P(m \mid A)$$

Burglary Example (revisited)



$\boldsymbol{P}(Burglary \mid JohnCalls = true, MaryCalls = true)$?

Hidden variables: $Y = \{Earthquake, Alarm\}$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{P}(B, j, m) = \alpha \underbrace{\sum_{E \in \{e, \neg e\}} \sum_{A \in \{a, \neg a\}}} \boldsymbol{P}(B, j, m, E, A)$$

// the textbook uses e, a instead of E, A in the summation.
// this is incorrect because e is a constant value standing
// for *Earthquake* = *true*. the random variable E can take
// on both e and $\neg e$ needed for the summation. you may
// also call it an abuse of the notation e and a.

evaluated for b (*Burglary* = *true*) based on

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i))$$

$$P(b \mid j, m) = \alpha \sum_E \sum_A P(b) P(E) P(A \mid b, E) P(j \mid A) P(m \mid A)$$

In the general case with n variables, there are 2^n summands, each as a product requires $O(n)$ computation time.

Expression Tree

Take advantage of the nested structure to move summations inwards as far as possible.

$$P(b \mid j,m) = \alpha \sum_{E} \sum_{A} P(b)P(E)P(A \mid b,E)P(j \mid A)P(m \mid A)$$

Expression Tree

Take advantage of the nested structure to move summations inwards as far as possible.

$$P(b \mid j, m) = \alpha \sum_{E} \sum_{A} P(b)P(E)P(A \mid b, E)P(j \mid A)P(m \mid A)$$

Expression Tree

Take advantage of the nested structure to move summations inwards as far as possible.

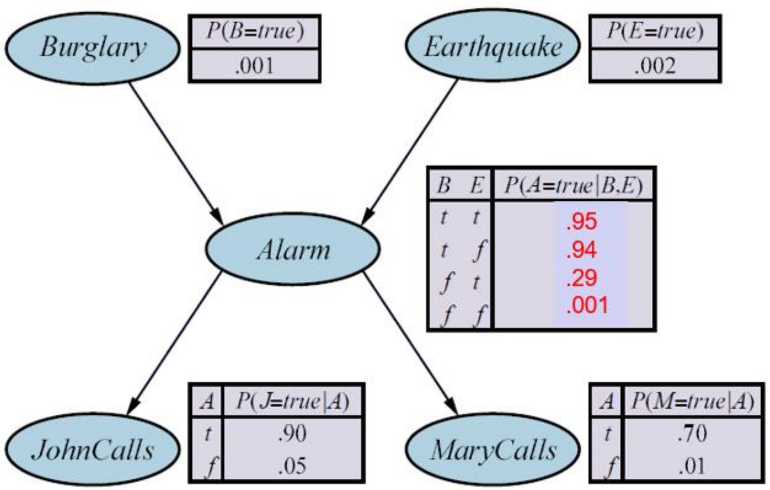
$$P(b \mid j, m) = \alpha \sum_{E} \sum_{A} P(b) P(E) P(A \mid b, E) P(j \mid A) P(m \mid A)$$

Expression Tree

$$P(b \mid j, m) = \alpha P(b) \sum_{E} P(E) \sum_{A} P(A \mid b, E) P(j \mid A) P(m \mid A)$$

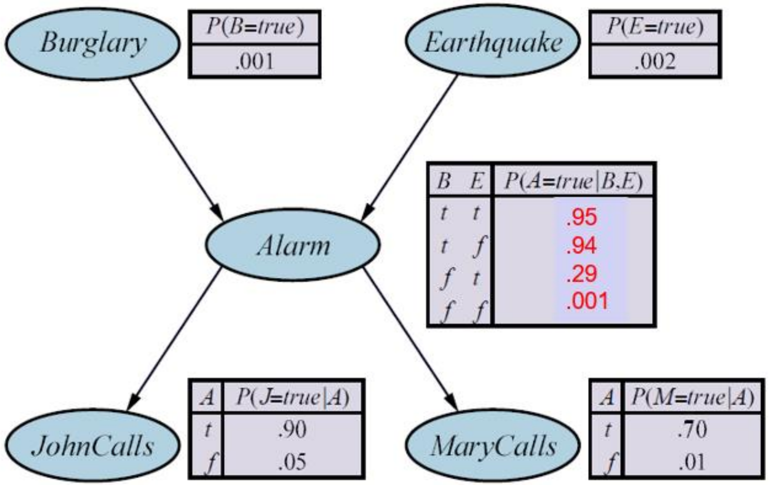
Expression Tree

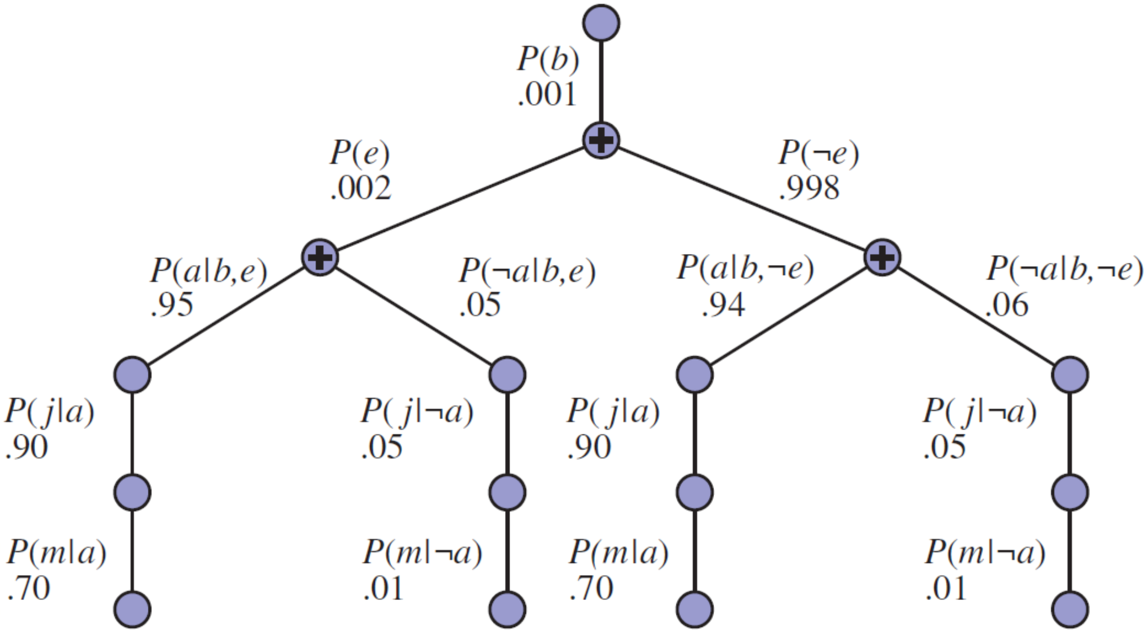
$$P(b \mid j, m) = \alpha P(b) \sum_{E} P(E) \sum_{A} P(A \mid b, E) P(j \mid A) P(m \mid A)$$



Expression Tree

$$P(b \mid j, m) = \alpha P(b) \sum_E P(E) \sum_A P(A \mid b, E) P(j \mid A) P(m \mid A)$$

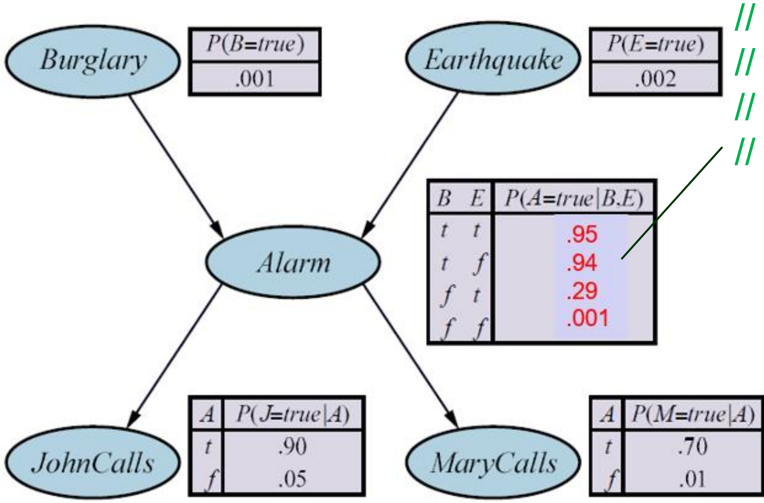


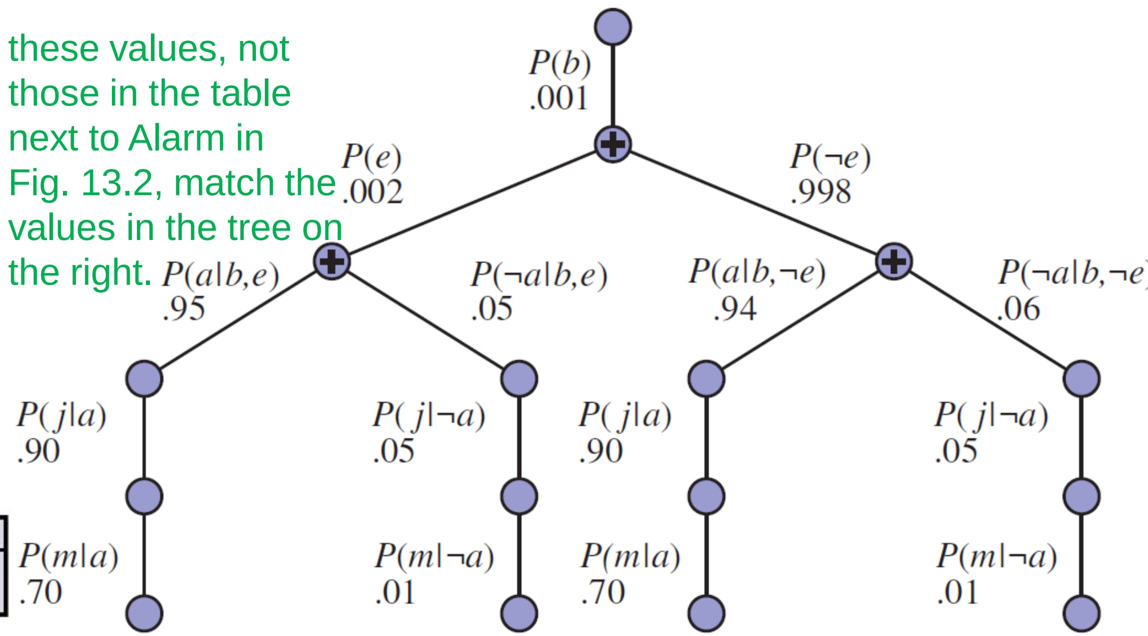


Expression Tree

$$P(b \mid j, m) = \alpha P(b) \sum_E P(E) \sum_A P(A \mid b, E) P(j \mid A) P(m \mid A)$$

// these values, not
// those in the table
// next to Alarm in
// Fig. 13.2, match the
// values in the tree on
// the right.

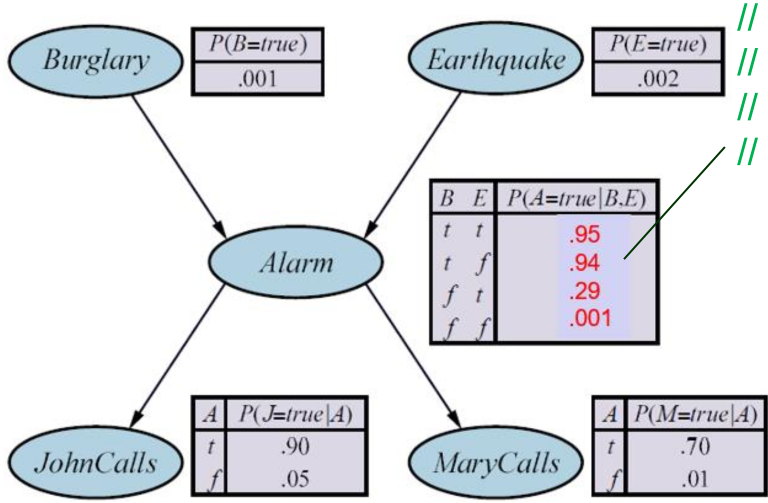


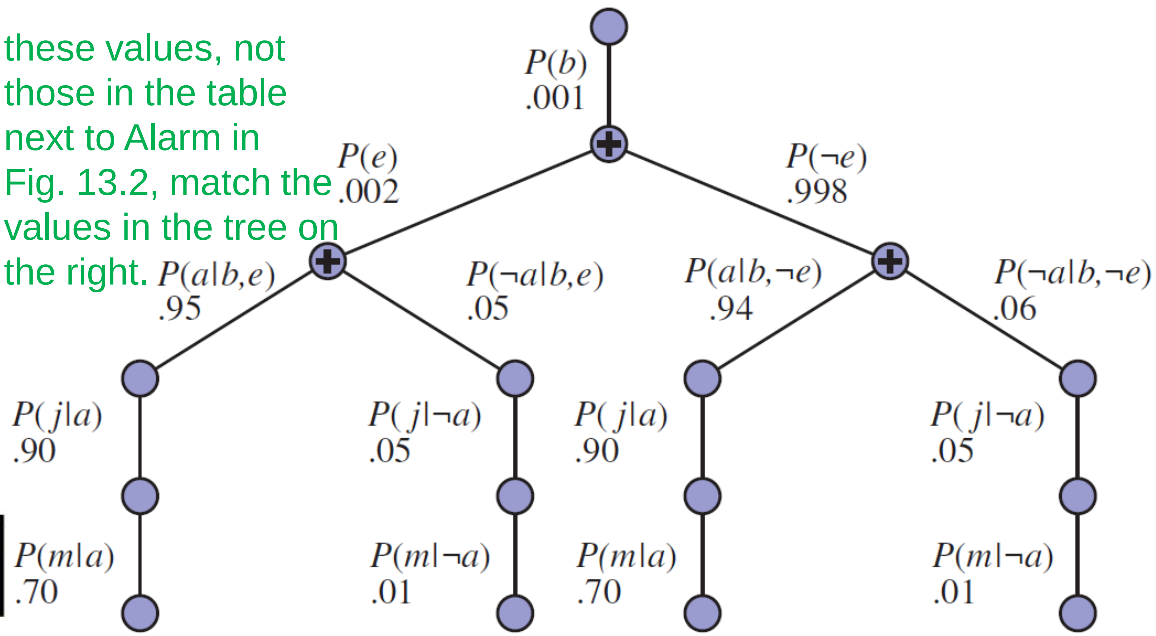


Expression Tree

$$P(b \mid j, m) = \alpha P(b) \sum_{E} P(E) \sum_{A} P(A \mid b, E) P(j \mid A) P(m \mid A)$$

// these values, not
// those in the table
// next to Alarm in
// Fig. 13.2, match the
// values in the tree on
// the right.



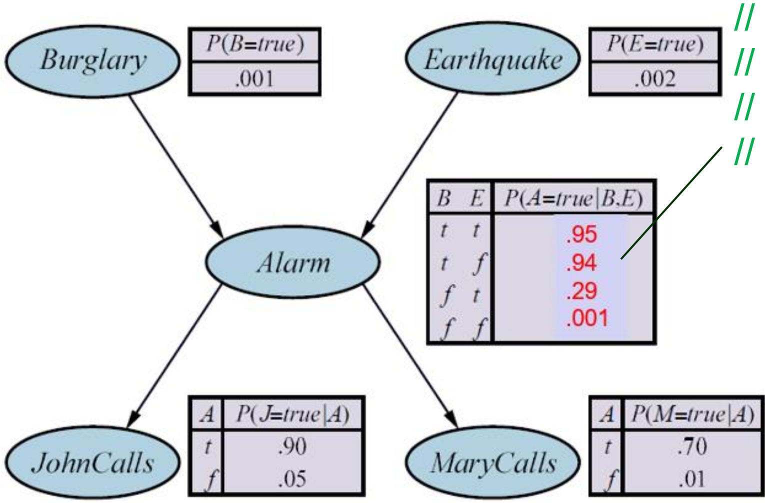


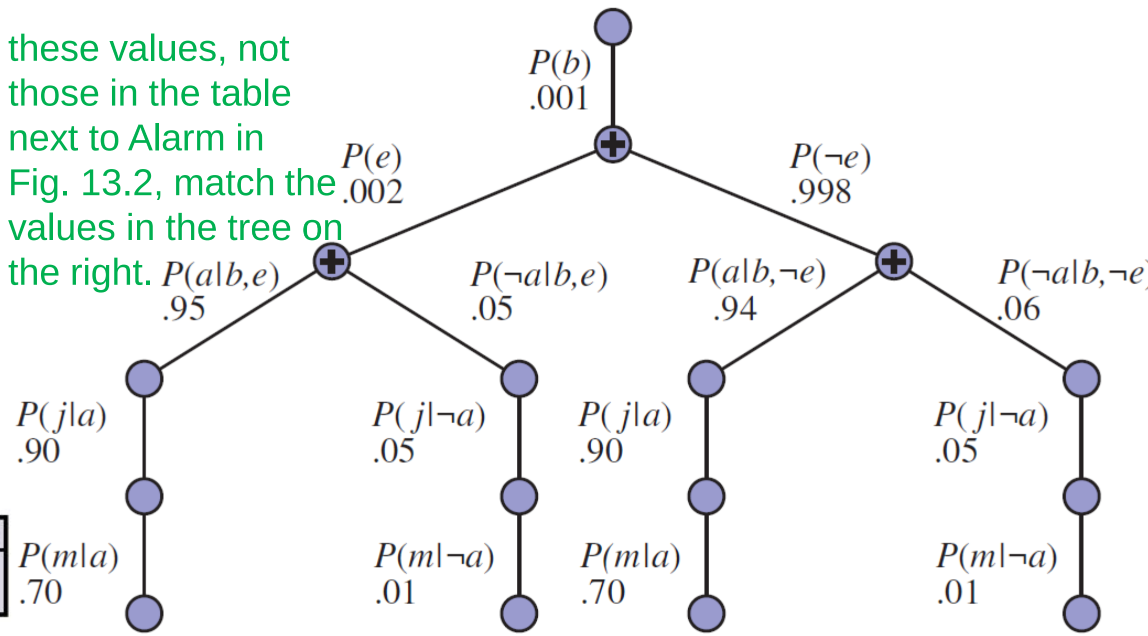
$$P(b \mid j, m) = \alpha \times 0.00059224$$

Expression Tree

$$P(b \mid j, m) = \alpha P(b) \sum_E P(E) \sum_A P(A \mid b, E) P(j \mid A) P(m \mid A)$$

// these values, not
// those in the table
// next to Alarm in
// Fig. 13.2, match the
// values in the tree on
// the right.



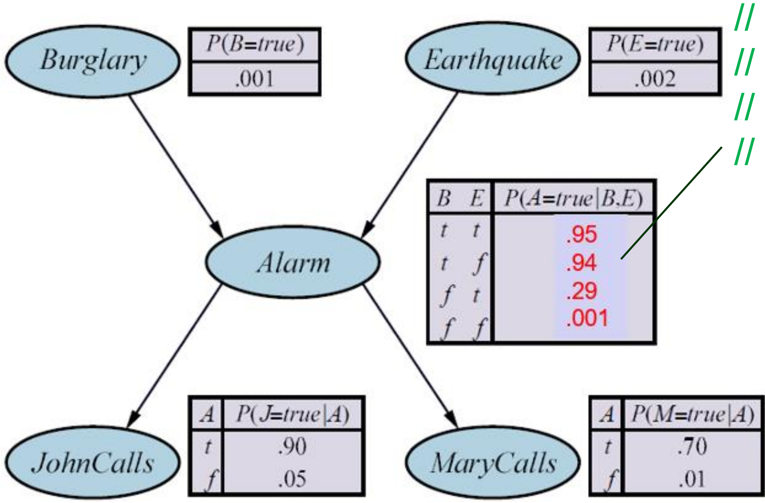


$$P(b \mid j, m) = \alpha \times 0.00059224$$

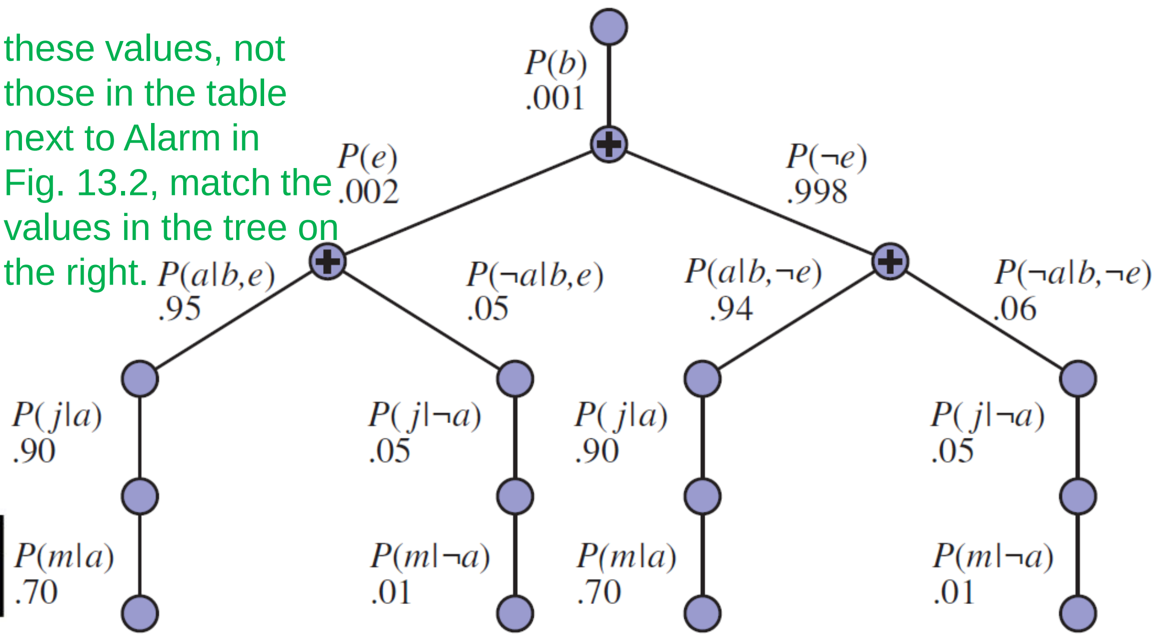
$$P(\neg b \mid j, m) = \alpha \times 0.0014919$$

Expression Tree

$$P(b \mid j, m) = \alpha P(b) \sum_{E} P(E) \sum_{A} P(A \mid b, E) P(j \mid A) P(m \mid A)$$



// these values, not
// those in the table
// next to Alarm in
// Fig. 13.2, match the
// values in the tree on
// the right.



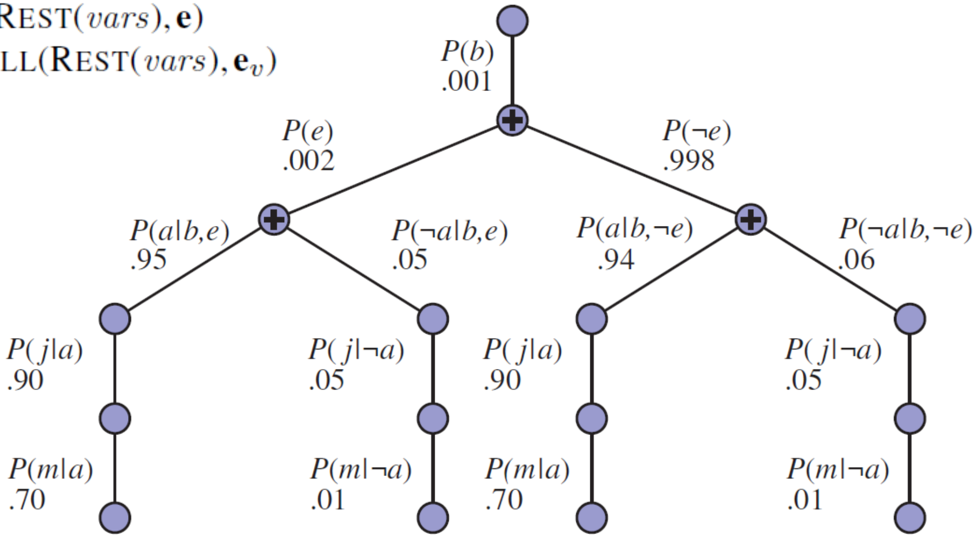
$$\left.\begin{aligned} P(b \mid j, m) &= \alpha \times 0.00059224 \\ P(\neg b \mid j, m) &= \alpha \times 0.0014919 \end{aligned}\right\} \Longrightarrow \mathbf{P}(B \mid j, m) = \alpha \langle 0.00059224, 0.0014919 \rangle \approx \langle 0.284, 0.716 \rangle$$

Enumeration Algorithm

function ENUMERATION-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 Bayes net with variables $vars$

  $\mathbf{Q}(X) \leftarrow$ a distribution over X, initially empty
  for each value x_i of X **do**
    $\mathbf{Q}(x_i) \leftarrow$ ENUMERATE-ALL($vars$, $\mathbf{e}_{x_i}$)
      where $\mathbf{e}_{x_i}$ is $\mathbf{e}$ extended with $X = x_i$
  return NORMALIZE($\mathbf{Q}(X)$)

function ENUMERATE-ALL($vars$, $\mathbf{e}$) **returns** a real number
  if EMPTY?($vars$) **then return** 1.0
  $V \leftarrow$ FIRST($vars$)
  if V is an evidence variable with value v in $\mathbf{e}$
    then return $P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}$)
    else return $\sum_v P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}_v$)
      where $\mathbf{e}_v$ is $\mathbf{e}$ extended with $V = v$



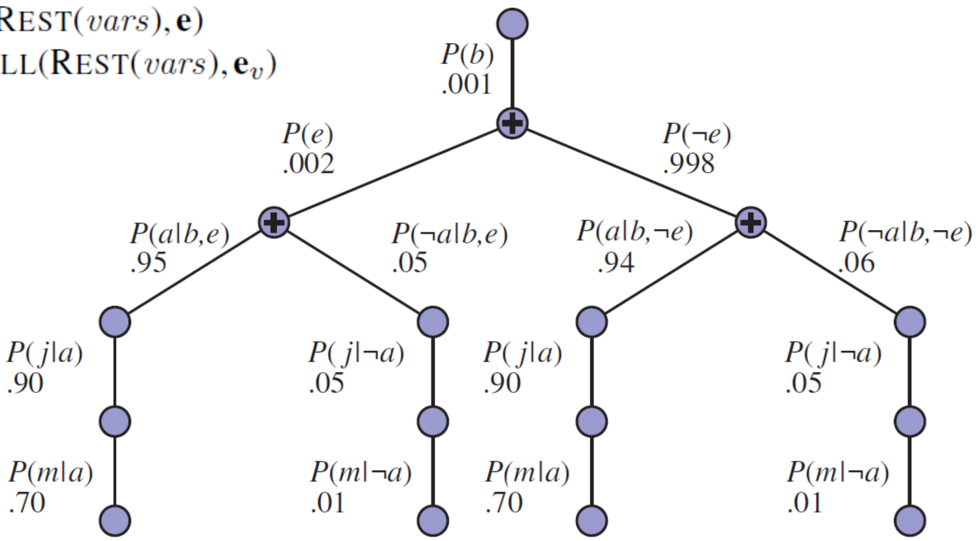
Enumeration Algorithm

function ENUMERATION-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 Bayes net with variables $vars$

  $\mathbf{Q}(X) \leftarrow$ a distribution over X, initially empty
  for each value x_i of X **do**
    $\mathbf{Q}(x_i) \leftarrow$ ENUMERATE-ALL($vars$, $\mathbf{e}_{x_i}$)
      where $\mathbf{e}_{x_i}$ is $\mathbf{e}$ extended with $X = x_i$
  return NORMALIZE($\mathbf{Q}(X)$)

function ENUMERATE-ALL($vars$, $\mathbf{e}$) **returns** a real number
  if EMPTY?($vars$) **then return** 1.0
  $V \leftarrow$ FIRST($vars$)
  if V is an evidence variable with value v in $\mathbf{e}$
    then return $P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}$)
    else return $\sum_v P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}_v$)
      where $\mathbf{e}_v$ is $\mathbf{e}$ extended with $V = v$

- Depth-first recursion



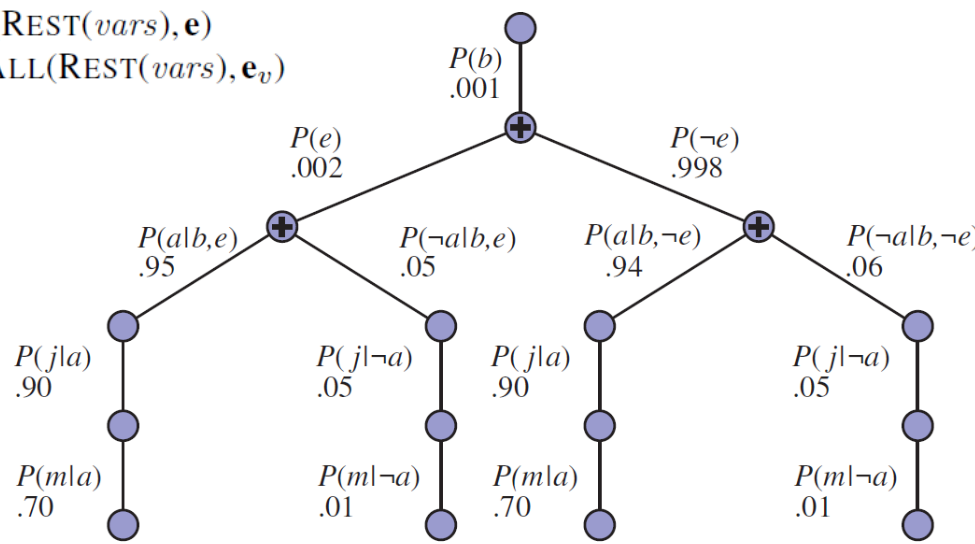
Enumeration Algorithm

function ENUMERATION-ASK(X, **e**, bn) **returns** a distribution over X
  inputs: X, the query variable
    e, observed values for variables **E**
    bn, a Bayes net with variables $vars$

  $\mathbf{Q}(X) \leftarrow$ a distribution over X, initially empty
  for each value x_i of X **do**
    $\mathbf{Q}(x_i) \leftarrow$ ENUMERATE-ALL($vars$, $\mathbf{e}_{x_i}$)
      where $\mathbf{e}_{x_i}$ is **e** extended with $X = x_i$
  return NORMALIZE($\mathbf{Q}(X)$)

function ENUMERATE-ALL($vars$, **e**) **returns** a real number
  if EMPTY?($vars$) **then return** 1.0
  $V \leftarrow$ FIRST($vars$)
  if V is an evidence variable with value v in **e**
    then return $P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), **e**)
    else return $\sum_v P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}_v$)
      where $\mathbf{e}_v$ is **e** extended with $V = v$

- Depth-first recursion
- $O(2^n)$ time for n variables



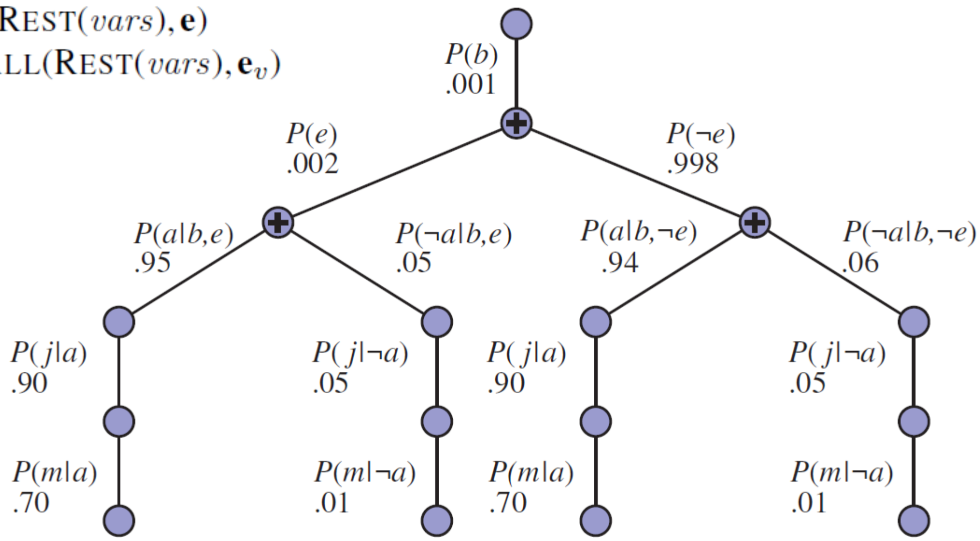
Enumeration Algorithm

function ENUMERATION-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 Bayes net with variables $vars$

  $\mathbf{Q}(X) \leftarrow$ a distribution over X, initially empty
  for each value x_i of X **do**
    $\mathbf{Q}(x_i) \leftarrow$ ENUMERATE-ALL($vars$, $\mathbf{e}_{x_i}$)
      where $\mathbf{e}_{x_i}$ is $\mathbf{e}$ extended with $X = x_i$
  return NORMALIZE($\mathbf{Q}(X)$)

function ENUMERATE-ALL($vars$, $\mathbf{e}$) **returns** a real number
  if EMPTY?($vars$) **then return** 1.0
  $V \leftarrow$ FIRST($vars$)
  if V is an evidence variable with value v in $\mathbf{e}$
    then return $P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}$)
    else return $\sum_v P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}_v$)
      where $\mathbf{e}_v$ is $\mathbf{e}$ extended with $V = v$

- Depth-first recursion
- $O(2^n)$ time for n variables
- Repeated evaluations of the same subexpressions



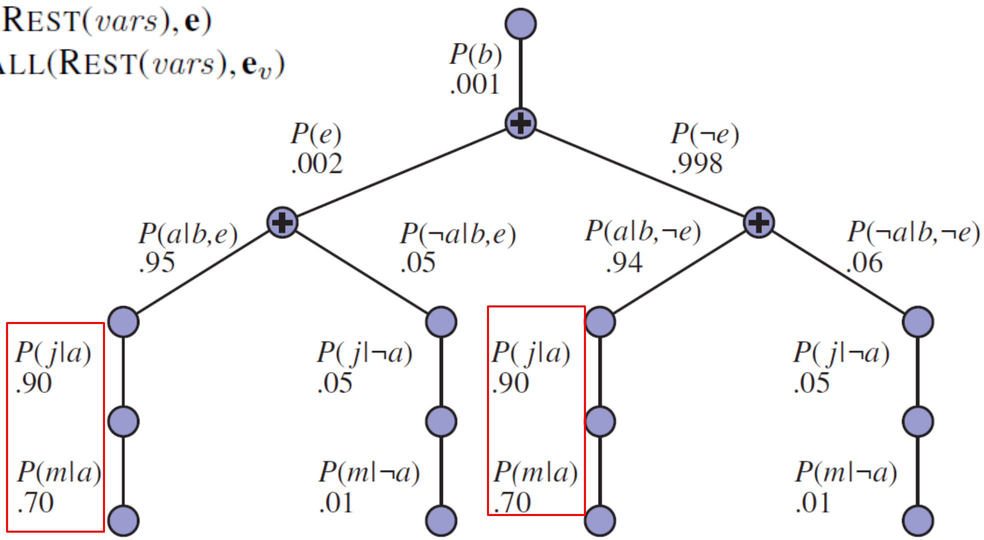
Enumeration Algorithm

function ENUMERATION-ASK(X, **e**, bn) **returns** a distribution over X
  inputs: X, the query variable
    e, observed values for variables **E**
    bn, a Bayes net with variables $vars$

  $\mathbf{Q}(X) \leftarrow$ a distribution over X, initially empty
  for each value x_i of X **do**
    $\mathbf{Q}(x_i) \leftarrow$ ENUMERATE-ALL($vars$, $\mathbf{e}_{x_i}$)
      where $\mathbf{e}_{x_i}$ is **e** extended with $X = x_i$
  return NORMALIZE($\mathbf{Q}(X)$)

function ENUMERATE-ALL($vars$, **e**) **returns** a real number
  if EMPTY?($vars$) **then return** 1.0
  $V \leftarrow$ FIRST($vars$)
  if V is an evidence variable with value v in **e**
    then return $P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), **e**)
    else return $\sum_v P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}_v$)
      where $\mathbf{e}_v$ is **e** extended with $V = v$

- Depth-first recursion
- $O(2^n)$ time for n variables
- Repeated evaluations of the same subexpressions



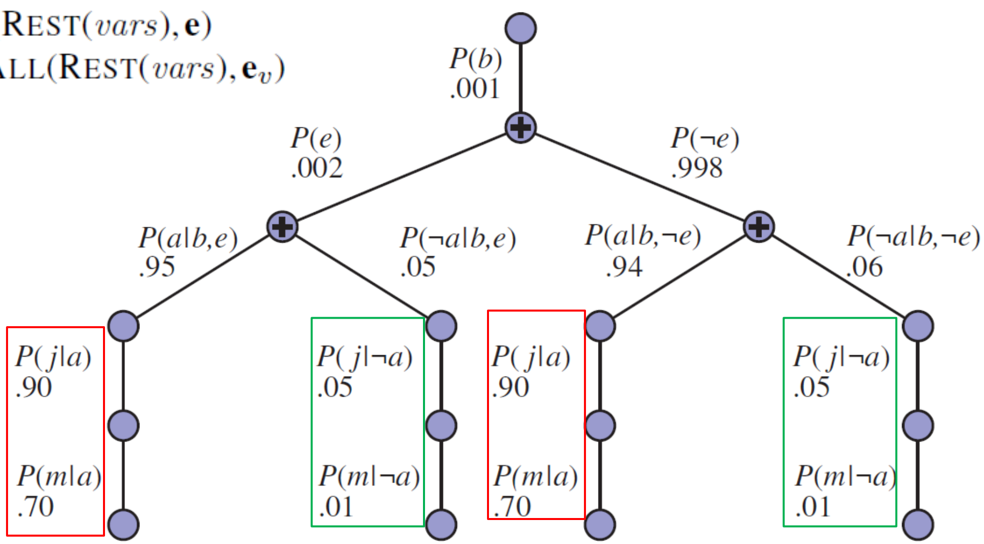
Enumeration Algorithm

function ENUMERATION-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 Bayes net with variables $vars$

  $\mathbf{Q}(X) \leftarrow$ a distribution over X, initially empty
  for each value x_i of X **do**
    $\mathbf{Q}(x_i) \leftarrow$ ENUMERATE-ALL($vars$, $\mathbf{e}_{x_i}$)
      where $\mathbf{e}_{x_i}$ is $\mathbf{e}$ extended with $X = x_i$
  return NORMALIZE($\mathbf{Q}(X)$)

function ENUMERATE-ALL($vars$, $\mathbf{e}$) **returns** a real number
  if EMPTY?($vars$) **then return** 1.0
  $V \leftarrow$ FIRST($vars$)
  if V is an evidence variable with value v in $\mathbf{e}$
    then return $P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}$)
    else return $\sum_v P(v \mid parents(V)) \times$ ENUMERATE-ALL(REST($vars$), $\mathbf{e}_v$)
      where $\mathbf{e}_v$ is $\mathbf{e}$ extended with $V = v$

- Depth-first recursion
- $O(2^n)$ time for n variables
- Repeated evaluations of the same subexpressions



II. Variable Elimination

Idea Do the calculation once and save the results for later use (like dynamic programming).

II. Variable Elimination

Idea Do the calculation once and save the results for later use (like dynamic programming).

- Evaluate an expression from right to left (i.e., bottom up in the expression tree).

$$P(b \mid j, m) = \alpha P(b) \sum_E P(E) \sum_A P(A \mid b, E) P(j \mid A) P(m \mid A)$$

II. Variable Elimination

Idea Do the calculation once and save the results for later use (like dynamic programming).

- Evaluate an expression from right to left (i.e., bottom up in the expression tree).

$$P(b \mid j, m) = \alpha P(b) \sum_E P(E) \sum_A P(A \mid b, E) P(j \mid A) P(m \mid A)$$

II. Variable Elimination

Idea Do the calculation once and save the results for later use (like dynamic programming).

- Evaluate an expression from right to left (i.e., bottom up in the expression tree).

$$P(b \mid j, m) = \alpha P(b) \sum_E P(E) \sum_A P(A \mid b, E) P(j \mid A) P(m \mid A)$$

$$\mathbf{P}(B \mid j, m) = \alpha \mathbf{P}(B) \sum_E P(E) \sum_A P(A \mid B, E) P(j \mid A) P(m \mid A)$$

II. Variable Elimination

Idea Do the calculation once and save the results for later use (like dynamic programming).

- Evaluate an expression from right to left (i.e., bottom up in the expression tree).

$$P(b \mid j, m) = \alpha P(b) \sum_E P(E) \sum_A P(A \mid b, E) P(j \mid A) P(m \mid A)$$

factors:

$$\mathbf{P}(B \mid j, m) = \alpha \underbrace{\mathbf{P}(B)}_{\boldsymbol{f}_1(B)} \sum_E \underbrace{P(E)}_{\boldsymbol{f}_2(E)} \sum_A \underbrace{P(A \mid B, E)}_{\boldsymbol{f}_3(A,B,E)} \underbrace{P(j \mid A)}_{\boldsymbol{f}_4(A)} \underbrace{P(m \mid A)}_{\boldsymbol{f}_5(A)}$$

II. Variable Elimination

Idea Do the calculation once and save the results for later use (like dynamic programming).

- Evaluate an expression from right to left (i.e., bottom up in the expression tree).

$$P(b \mid j, m) = \alpha P(b) \sum_E P(E) \sum_A P(A \mid b, E) P(j \mid A) P(m \mid A)$$

$$\mathbf{P}(B \mid j, m) = \alpha \underbrace{\mathbf{P}(B)}_{} \sum_E \underbrace{P(E)}_{} \sum_A \underbrace{P(A \mid B, E)}_{} \underbrace{P(j \mid A)}_{} \underbrace{P(m \mid A)}_{}$$

factors:	$\boldsymbol{f}_1(B)$	$\boldsymbol{f}_2(E)$	$\boldsymbol{f}_3(A, B, E)$	$\boldsymbol{f}_4(A)$	$\boldsymbol{f}_5(A)$
dimensions:	2×1	2×1	$2 \times 2 \times 2$	2×1	2×1

Evaluation Example

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_{E} \boldsymbol{f_2}(E) \times \sum_{A} \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A)$$

Evaluation Example

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_E \boldsymbol{f_2}(E) \times \sum_A \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A)$$

- Sum out A from the product of $\boldsymbol{f}_3, \boldsymbol{f}_4, \boldsymbol{f}_5$.

$$\underset{2 \times 2}{\boldsymbol{f}_6(B, E)} = \sum_{A \in \{a, \neg a\}} \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A)$$

Evaluation Example

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_E \boldsymbol{f_2}(E) \times \sum_A \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A)$$

- Sum out A from the product of $\boldsymbol{f}_3, \boldsymbol{f}_4, \boldsymbol{f}_5$.

$$\underset{2 \times 2}{\boldsymbol{f}_6(B, E)} = \sum_{A \in \{a, \neg a\}} \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A)$$

$$= \left(\underset{2 \times 2}{\boldsymbol{f}_3(a, B, E)} \times P(a) \times P(a)\right) + \left(\underset{2 \times 2}{\boldsymbol{f}_3(\neg a, B, E)} \times P(\neg a) \times P(\neg a)\right)$$

Evaluation Example

$$\boldsymbol{P}(B \mid j,m) = \alpha \boldsymbol{f_1}(B) \times \sum_E \boldsymbol{f_2}(E) \times \sum_A \boldsymbol{f_3}(A,B,E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A)$$

- Sum out A from the product of $\boldsymbol{f}_3, \boldsymbol{f}_4, \boldsymbol{f}_5$.

$$\underset{2\times 2}{\boldsymbol{f}_6(B,E)} = \sum_{A\in\{a,\neg a\}} \boldsymbol{f_3}(A,B,E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A)$$

$$= \left(\underset{2\times 2}{\boldsymbol{f}_3(a,B,E)} \times P(a) \times P(a)\right) + \left(\underset{2\times 2}{\boldsymbol{f}_3(\neg a,B,E)} \times P(\neg a) \times P(\neg a)\right)$$

$$\boldsymbol{P}(B \mid j,m) = \alpha \boldsymbol{f_1}(B) \times \sum_E \boldsymbol{f}_2(E) \times \boldsymbol{f}_6(B,E)$$

Evaluation Example

$$\boldsymbol{P}(B \mid j,m) = \alpha \boldsymbol{f}_{\mathbf{1}}(B) \times \sum_{E} \boldsymbol{f}_{\mathbf{2}}(E) \times \sum_{A} \boldsymbol{f}_{\mathbf{3}}(A,B,E) \times \boldsymbol{f}_{\mathbf{4}}(A) \times \boldsymbol{f}_{\mathbf{5}}(A)$$

- Sum out A from the product of $\boldsymbol{f}_3, \boldsymbol{f}_4, \boldsymbol{f}_5$.

$$\underset{2\times 2}{\boldsymbol{f}_6(B,E)} = \sum_{A\in\{a,\neg a\}} \boldsymbol{f}_{\mathbf{3}}(A,B,E) \times \boldsymbol{f}_{\mathbf{4}}(A) \times \boldsymbol{f}_{\mathbf{5}}(A)$$

$$= \left(\underset{2\times 2}{\boldsymbol{f}_3(a,B,E)} \times P(a) \times P(a)\right) + \left(\underset{2\times 2}{\boldsymbol{f}_3(\neg a,B,E)} \times P(\neg a) \times P(\neg a)\right)$$

$$\boldsymbol{P}(B \mid j,m) = \alpha \boldsymbol{f}_{\mathbf{1}}(B) \times \sum_{E} \boldsymbol{f}_2(E) \times \boldsymbol{f}_6(B,E)$$

- Sum out E from the product of $\boldsymbol{f}_2$ and $\boldsymbol{f}_6$.

$$\boldsymbol{f}_7(B) = \sum_{E\in\{e,\neg e\}} \boldsymbol{f}_{\mathbf{2}}(E) \times \boldsymbol{f}_{\mathbf{6}}(B,E) = P(e) \times \underset{2\times 1}{\boldsymbol{f}_6(B,e)} + P(\neg e) \times \underset{2\times 1}{\boldsymbol{f}_6(B,\neg e)}$$

Evaluation Example

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_E \boldsymbol{f_2}(E) \times \sum_A \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A)$$

- Sum out A from the product of $\boldsymbol{f}_3, \boldsymbol{f}_4, \boldsymbol{f}_5$.

$$\underset{2 \times 2}{\boldsymbol{f}_6(B, E)} = \sum_{A \in \{a, \neg a\}} \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A)$$

$$= \left(\underset{2 \times 2}{\boldsymbol{f}_3(a, B, E)} \times P(a) \times P(a)\right) + \left(\underset{2 \times 2}{\boldsymbol{f}_3(\neg a, B, E)} \times P(\neg a) \times P(\neg a)\right)$$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_E \boldsymbol{f}_2(E) \times \boldsymbol{f}_6(B, E)$$

- Sum out E from the product of $\boldsymbol{f}_2$ and $\boldsymbol{f}_6$.

$$\boldsymbol{f}_7(B) = \sum_{E \in \{e, \neg e\}} \boldsymbol{f_2}(E) \times \boldsymbol{f_6}(B, E) = P(e) \times \underset{2 \times 1}{\boldsymbol{f}_6(B, e)} + P(\neg e) \times \underset{2 \times 1}{\boldsymbol{f}_6(B, \neg e)}$$

- Finally, carry out the following pointwise product:

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \boldsymbol{f_7}(B)$$

Pointwise Product of Two Factors

$$\boldsymbol{f}(X_1, \dots, X_j, Y_1, \dots, Y_k) \times \boldsymbol{g}(Y_1, \dots, Y_k, Z_1, \dots, Z_l) = \boldsymbol{h}(X_1, \dots, X_j, Y_1, \dots, Y_k, Z_1, \dots, Z_l)$$

X	Y	$\mathbf{f}(X,Y)$	Y	Z	$\mathbf{g}(Y,Z)$	X	Y	Z	$\mathbf{h}(X,Y,Z)$
t	t	.3	t	t	.2	t	t	t	$.3 \times .2 = .06$
t	f	.7	t	f	.8	t	t	f	$.3 \times .8 = .24$
f	t	.9	f	t	.6	t	f	t	$.7 \times .6 = .42$
f	f	.1	f	f	.4	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Pointwise Product of Two Factors

$$\boldsymbol{f}(X_1, \ldots, X_j, Y_1, \ldots, Y_k) \times \boldsymbol{g}(Y_1, \ldots, Y_k, Z_1, \ldots, Z_l) = \boldsymbol{h}(X_1, \ldots, X_j, Y_1, \ldots, Y_k, Z_1, \ldots, Z_l)$$

common variables

X	Y	$\mathbf{f}(X,Y)$	Y	Z	$\mathbf{g}(Y,Z)$	X	Y	Z	$\mathbf{h}(X,Y,Z)$
t	t	.3	t	t	.2	t	t	t	$.3 \times .2 = .06$
t	f	.7	t	f	.8	t	t	f	$.3 \times .8 = .24$
f	t	.9	f	t	.6	t	f	t	$.7 \times .6 = .42$
f	f	.1	f	f	.4	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Pointwise Product of Two Factors

$$\boldsymbol{f}(X_1,\dots,X_j,\underbrace{Y_1,\dots,Y_k}_{\text{common variables}}) \times \boldsymbol{g}(\underbrace{Y_1,\dots,Y_k}_{\text{common variables}},Z_1,\dots,Z_l) = \boldsymbol{h}(X_1,\dots,X_j,Y_1,\dots,Y_k,Z_1,\dots,Z_l)$$

X	Y	$\mathbf{f}(X,Y)$	Y	Z	$\mathbf{g}(Y,Z)$	X	Y	Z	$\mathbf{h}(X,Y,Z)$
t	t	.3	t	t	.2	t	t	t	$.3 \times .2 = .06$
t	f	.7	t	f	.8	t	t	f	$.3 \times .8 = .24$
f	t	.9	f	t	.6	t	f	t	$.7 \times .6 = .42$
f	f	.1	f	f	.4	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Pointwise Product of Two Factors

$$\boldsymbol{f}(X_1, \ldots, X_j, Y_1, \ldots, Y_k) \times \boldsymbol{g}(Y_1, \ldots, Y_k, Z_1, \ldots, Z_l) = \boldsymbol{h}(X_1, \ldots, X_j, Y_1, \ldots, Y_k, Z_1, \ldots, Z_l)$$

common variables

X	Y	$\mathbf{f}(X, Y)$	Y	Z	$\mathbf{g}(Y, Z)$	X	Y	Z	$\mathbf{h}(X, Y, Z)$
t	t	.3	t	t	.2	t	t	t	$.3 \times .2 = .06$
t	f	.7	t	f	.8	t	t	f	$.3 \times .8 = .24$
f	t	.9	f	t	.6	t	f	t	$.7 \times .6 = .42$
f	f	.1	f	f	.4	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Pointwise Product of Two Factors

$$\boldsymbol{f}(X_1, \ldots, X_j, \underbrace{Y_1, \ldots, Y_k}_{\text{common variables}}) \times \boldsymbol{g}(\underbrace{Y_1, \ldots, Y_k}, Z_1, \ldots, Z_l) = \boldsymbol{h}(X_1, \ldots, X_j, Y_1, \ldots, Y_k, Z_1, \ldots, Z_l)$$

X	Y	$\mathbf{f}(X, Y)$	Y	Z	$\mathbf{g}(Y, Z)$	X	Y	Z	$\mathbf{h}(X, Y, Z)$
t	t	.3	t	t	.2	t	t	t	$.3 \times .2 = .06$
t	f	.7	t	f	.8	t	t	f	$.3 \times .8 = .24$
f	t	.9	f	t	.6	t	f	t	$.7 \times .6 = .42$
f	f	.1	f	f	.4	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Summing Out a Variable

This operation over a product of factors is carried out by adding up the submatrices, each for one value of the same variable.

$$\boldsymbol{l}(Y,Z) = \sum_X \boldsymbol{h}(X,Y,Z) = \boldsymbol{h}(x,Y,Z) + \boldsymbol{h}(\neg x,Y,Z)$$

X	Y	$\mathbf{f}(X,Y)$	Y	Z	$\mathbf{g}(Y,Z)$	X	Y	Z	$\mathbf{h}(X,Y,Z)$
t	t	.3	t	t	.2	t	t	t	$.3 \times .2 = .06$
t	f	.7	t	f	.8	t	t	f	$.3 \times .8 = .24$
f	t	.9	f	t	.6	t	f	t	$.7 \times .6 = .42$
f	f	.1	f	f	.4	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Summing Out a Variable

This operation over a product of factors is carried out by adding up the submatrices, each for one value of the same variable.

$$\boldsymbol{l}(Y,Z) = \sum_{X} \boldsymbol{h}(X,Y,Z) = \boldsymbol{h}(x,Y,Z) + \boldsymbol{h}(\neg x,Y,Z)$$

X	Y	$\mathbf{f}(X,Y)$	Y	Z	$\mathbf{g}(Y,Z)$	X	Y	Z	$\mathbf{h}(X,Y,Z)$
t	t	.3	t	t	.2	t	t	t	$.3 \times .2 = .06$
t	f	.7	t	f	.8	t	t	f	$.3 \times .8 = .24$
f	t	.9	f	t	.6	t	f	t	$.7 \times .6 = .42$
f	f	.1	f	f	.4	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Summing Out a Variable

This operation over a product of factors is carried out by adding up the submatrices, each for one value of the same variable.

$$\boldsymbol{l}(Y,Z) = \sum_X \boldsymbol{h}(X,Y,Z) = \boldsymbol{h}(x,Y,Z) + \boldsymbol{h}(\neg x,Y,Z)$$

X	Y	$\mathbf{f}(X,Y)$	Y	Z	$\mathbf{g}(Y,Z)$	X	Y	Z	$\mathbf{h}(X,Y,Z)$
t	t	$.3$	t	t	$.2$	t	t	t	$.3 \times .2 = .06$
t	f	$.7$	t	f	$.8$	t	t	f	$.3 \times .8 = .24$
f	t	$.9$	f	t	$.6$	t	f	t	$.7 \times .6 = .42$
f	f	$.1$	f	f	$.4$	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Summing Out a Variable

This operation over a product of factors is carried out by adding up the submatrices, each for one value of the same variable.

$$\boldsymbol{l}(Y,Z) = \sum_X \boldsymbol{h}(X,Y,Z) = \boldsymbol{h}(x,Y,Z) + \boldsymbol{h}(\neg x,Y,Z)$$

$$= \begin{pmatrix} .06 & .24 \\ .42 & .28 \end{pmatrix}$$

X	Y	$\mathbf{f}(X,Y)$	Y	Z	$\mathbf{g}(Y,Z)$	X	Y	Z	$\mathbf{h}(X,Y,Z)$
t	t	.3	t	t	.2	t	t	t	$.3 \times .2 = .06$
t	f	.7	t	f	.8	t	t	f	$.3 \times .8 = .24$
f	t	.9	f	t	.6	t	f	t	$.7 \times .6 = .42$
f	f	.1	f	f	.4	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Summing Out a Variable

This operation over a product of factors is carried out by adding up the submatrices, each for one value of the same variable.

$$\boldsymbol{l}(Y,Z) = \sum_X \boldsymbol{h}(X,Y,Z) = \boldsymbol{h}(x,Y,Z) + \boldsymbol{h}(\neg x,Y,Z)$$

$$= \begin{pmatrix} .06 & .24 \\ .42 & .28 \end{pmatrix}$$

X	Y	$\mathbf{f}(X,Y)$	Y	Z	$\mathbf{g}(Y,Z)$	X	Y	Z	$\mathbf{h}(X,Y,Z)$
t	t	$.3$	t	t	$.2$	t	t	t	$.3 \times .2 = .06$
t	f	$.7$	t	f	$.8$	t	t	f	$.3 \times .8 = .24$
f	t	$.9$	f	t	$.6$	t	f	t	$.7 \times .6 = .42$
f	f	$.1$	f	f	$.4$	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Summing Out a Variable

This operation over a product of factors is carried out by adding up the submatrices, each for one value of the same variable.

$$\boldsymbol{l}(Y,Z) = \sum_X \boldsymbol{h}(X,Y,Z) = \boldsymbol{h}(x,Y,Z) + \boldsymbol{h}(\neg x,Y,Z)$$

$$= \begin{pmatrix} .06 & .24 \\ .42 & .28 \end{pmatrix}$$

X	Y	$\mathbf{f}(X,Y)$	Y	Z	$\mathbf{g}(Y,Z)$	X	Y	Z	$\mathbf{h}(X,Y,Z)$
t	t	$.3$	t	t	$.2$	t	t	t	$.3 \times .2 = .06$
t	f	$.7$	t	f	$.8$	t	t	f	$.3 \times .8 = .24$
f	t	$.9$	f	t	$.6$	t	f	t	$.7 \times .6 = .42$
f	f	$.1$	f	f	$.4$	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Summing Out a Variable

This operation over a product of factors is carried out by adding up the submatrices, each for one value of the same variable.

$$\boldsymbol{l}(Y,Z) = \sum_X \boldsymbol{h}(X,Y,Z) = \boldsymbol{h}(x,Y,Z) + \boldsymbol{h}(\neg x,Y,Z)$$

$$= \begin{pmatrix} .06 & .24 \\ .42 & .28 \end{pmatrix} + \begin{pmatrix} .18 & .72 \\ .06 & .04 \end{pmatrix}$$

X	Y	$\mathbf{f}(X,Y)$	Y	Z	$\mathbf{g}(Y,Z)$	X	Y	Z	$\mathbf{h}(X,Y,Z)$
t	t	$.3$	t	t	$.2$	t	t	t	$.3 \times .2 = .06$
t	f	$.7$	t	f	$.8$	t	t	f	$.3 \times .8 = .24$
f	t	$.9$	f	t	$.6$	t	f	t	$.7 \times .6 = .42$
f	f	$.1$	f	f	$.4$	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Summing Out a Variable

This operation over a product of factors is carried out by adding up the submatrices, each for one value of the same variable.

$$\boldsymbol{l}(Y,Z) = \sum_X \boldsymbol{h}(X,Y,Z) = \boldsymbol{h}(x,Y,Z) + \boldsymbol{h}(\neg x,Y,Z)$$

$$= \begin{pmatrix} .06 & .24 \\ .42 & .28 \end{pmatrix} + \begin{pmatrix} .18 & .72 \\ .06 & .04 \end{pmatrix} = \begin{pmatrix} .24 & .96 \\ .48 & .32 \end{pmatrix}$$

X	Y	$\mathbf{f}(X,Y)$	Y	Z	$\mathbf{g}(Y,Z)$	X	Y	Z	$\mathbf{h}(X,Y,Z)$
t	t	$.3$	t	t	$.2$	t	t	t	$.3 \times .2 = .06$
t	f	$.7$	t	f	$.8$	t	t	f	$.3 \times .8 = .24$
f	t	$.9$	f	t	$.6$	t	f	t	$.7 \times .6 = .42$
f	f	$.1$	f	f	$.4$	t	f	f	$.7 \times .4 = .28$
						f	t	t	$.9 \times .2 = .18$
						f	t	f	$.9 \times .8 = .72$
						f	f	t	$.1 \times .6 = .06$
						f	f	f	$.1 \times .4 = .04$

Variable Elimination Algorithm

Move outside the summation any factor independent of the variable to be summed out.

Variable Elimination Algorithm

Move outside the summation any factor independent of the variable to be summed out.

$$\sum_X \boldsymbol{f}(X,Y) \times \boldsymbol{g}(Y,Z) = \boldsymbol{g}(Y,Z) \times \sum_X \boldsymbol{f}(X,Y)$$

Variable Elimination Algorithm

Move outside the summation any factor independent of the variable to be summed out.

$$\sum_X \boldsymbol{f}(X,Y) \times \boldsymbol{g}(Y,Z) = \boldsymbol{g}(Y,Z) \times \sum_X \boldsymbol{f}(X,Y)$$

function ELIMINATION-ASK(X, **e**, bn) **returns** a distribution over X
  inputs: X, the query variable
    e, observed values for variables **E**
    bn, a Bayesian network with variables $vars$

$factors \leftarrow [\,]$
for each V **in** ORDER($vars$) **do**
  $factors \leftarrow$ [MAKE-FACTOR(V, **e**)] + $factors$
  if V is a hidden variable **then** $factors \leftarrow$ SUM-OUT(V, $factors$)
return NORMALIZE(POINTWISE-PRODUCT($factors$))

Variable Elimination Algorithm

Move outside the summation any factor independent of the variable to be summed out.

$$\sum_{X} \boldsymbol{f}(X,Y) \times \boldsymbol{g}(Y,Z) = \boldsymbol{g}(Y,Z) \times \sum_{X} \boldsymbol{f}(X,Y)$$

function ELIMINATION-ASK(X, **e**, bn) **returns** a distribution over X
  inputs: X, the query variable
      e, observed values for variables **E**
      bn, a Bayesian network with variables $vars$

$factors \leftarrow [\,]$?
for each V **in** $\boxed{\text{ORDER}}$($vars$) **do**
  $factors \leftarrow$ [MAKE-FACTOR(V, **e**)] $+$ $factors$
  if V is a hidden variable **then** $factors \leftarrow$ SUM-OUT(V, $factors$)
return NORMALIZE(POINTWISE-PRODUCT($factors$))

III. Ordering Variables to Eliminate

$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i))$$

III. Ordering Variables to Eliminate

- Every choice of ordering yields a valid algorithm. ⇦ $P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid \mathit{Parents}(X_i))$

III. Ordering Variables to Eliminate

- Every choice of ordering yields a valid algorithm. ⇐ $P(x_1, \dots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i))$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_E \boldsymbol{f_2}(E) \times \sum_A \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A)$$

(Order: A, E)

III. Ordering Variables to Eliminate

- Every choice of ordering yields a valid algorithm. ⇦ $P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i))$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_E \boldsymbol{f_2}(E) \times \sum_A \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A) \qquad \text{(Order: } A, E\text{)}$$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_A \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A) \times \sum_E \boldsymbol{f_2}(E) \times \boldsymbol{f_3}(A, B, E) \qquad \text{(Order: } E, A\text{)}$$

III. Ordering Variables to Eliminate

- Every choice of ordering yields a valid algorithm. ⇐ $P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i))$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f}_{\mathbf{1}}(B) \times \sum_{E} \boldsymbol{f}_{\mathbf{2}}(E) \times \sum_{A} \boldsymbol{f}_{\mathbf{3}}(A, B, E) \times \boldsymbol{f}_{\mathbf{4}}(A) \times \boldsymbol{f}_{\mathbf{5}}(A) \qquad \text{(Order: } A, E\text{)}$$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f}_{\mathbf{1}}(B) \times \sum_{A} \boldsymbol{f}_{\mathbf{4}}(A) \times \boldsymbol{f}_{\mathbf{5}}(A) \times \sum_{E} \boldsymbol{f}_{\mathbf{2}}(E) \times \boldsymbol{f}_{\mathbf{3}}(A, B, E) \qquad \text{(Order: } E, A\text{)}$$

- Different orderings generates different intermediate factors.

III. Ordering Variables to Eliminate

- Every choice of ordering yields a valid algorithm. ⇦ $P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i))$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_E \boldsymbol{f_2}(E) \times \sum_A \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A) \qquad \text{(Order: } A, E\text{)}$$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_A \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A) \times \sum_E \boldsymbol{f_2}(E) \times \boldsymbol{f_3}(A, B, E) \qquad \text{(Order: } E, A\text{)}$$

- Different orderings generates different intermediate factors.

- Time and space are dominated by the size of the largest factor constructed, which is subject to two factors:

III. Ordering Variables to Eliminate

- Every choice of ordering yields a valid algorithm. ⇦ $P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i))$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_{E} \boldsymbol{f_2}(E) \times \sum_{A} \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A) \qquad \text{(Order: } A, E\text{)}$$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_{A} \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A) \times \sum_{E} \boldsymbol{f_2}(E) \times \boldsymbol{f_3}(A, B, E) \qquad \text{(Order: } E, A\text{)}$$

- Different orderings generates different intermediate factors.

- Time and space are dominated by the size of the largest factor constructed, which is subject to two factors:

  - ordering of variables
  - structure of the network

III. Ordering Variables to Eliminate

- Every choice of ordering yields a valid algorithm. $\Leftarrow$ $P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i))$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_E \boldsymbol{f_2}(E) \times \sum_A \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A) \qquad \text{(Order: } A, E\text{)}$$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_A \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A) \times \sum_E \boldsymbol{f_2}(E) \times \boldsymbol{f_3}(A, B, E) \qquad \text{(Order: } E, A\text{)}$$

- Different orderings generates different intermediate factors.

- Time and space are dominated by the size of the largest factor constructed, which is subject to two factors:
  - ordering of variables
  - structure of the network

- It is intractable to determine the optimal order.

III. Ordering Variables to Eliminate

- Every choice of ordering yields a valid algorithm. ⇦ $P(x_1, \dots, x_n) = \prod_{i=1}^{n} P(x_i \mid Parents(X_i))$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_E \boldsymbol{f_2}(E) \times \sum_A \boldsymbol{f_3}(A, B, E) \times \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A) \qquad \text{(Order: } A, E\text{)}$$

$$\boldsymbol{P}(B \mid j, m) = \alpha \boldsymbol{f_1}(B) \times \sum_A \boldsymbol{f_4}(A) \times \boldsymbol{f_5}(A) \times \sum_E \boldsymbol{f_2}(E) \times \boldsymbol{f_3}(A, B, E) \qquad \text{(Order: } E, A\text{)}$$

- Different orderings generates different intermediate factors.

- Time and space are dominated by the size of the largest factor constructed, which is subject to two factors:
  - ordering of variables
  - structure of the network

- It is intractable to determine the optimal order.

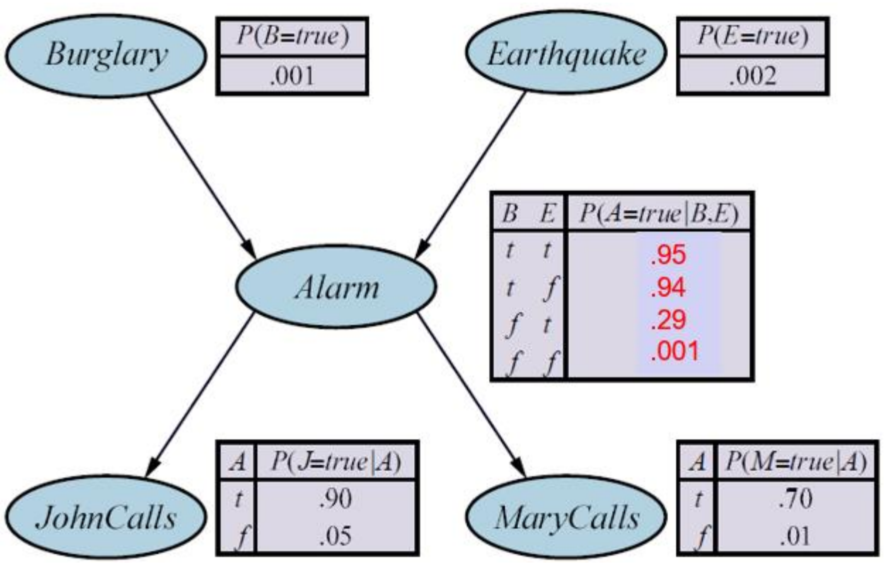
- Use a greedy heuristic: eliminate whichever variable minimizes the size of the next factor to be constructed.

Irrelevant Variable

$\boldsymbol{P}(JohnCalls \mid Burglary = true)$?

Irrelevant Variable

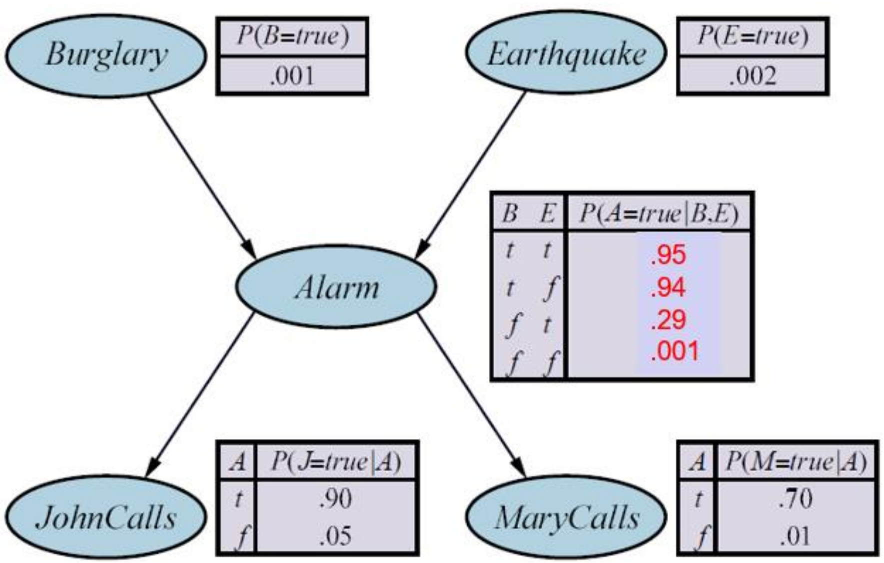
$\boldsymbol{P}(JohnCalls \mid Burglary = true)$?



Irrelevant Variable

$\mathbf{P}(JohnCalls \mid Burglary = true)$?

$$\mathbf{P}(J \mid b) = \alpha P(b) \sum_E P(E) \sum_A P(A \mid b, E)\, \mathbf{P}(J \mid A) \sum_M P(M \mid A)$$

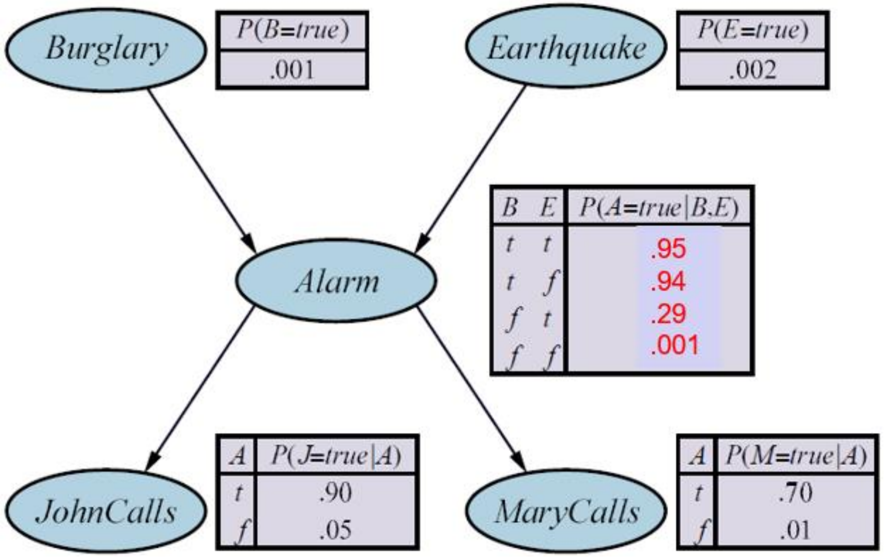


Irrelevant Variable

$\boldsymbol{P}(JohnCalls \mid Burglary = true)?$

$$\boldsymbol{P}(J \mid b) = \alpha P(b) \sum_{E} P(E) \sum_{A} P(A \mid b, E)\, \boldsymbol{P}(J \mid A) \sum_{M} P(M \mid A)$$

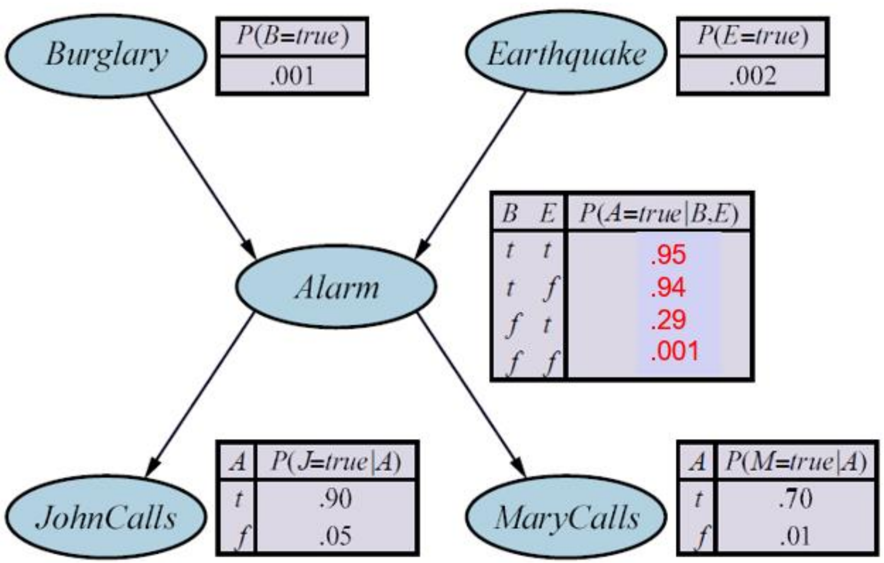
$$\sum_{M} P(M \mid A) = 1$$



Irrelevant Variable

$\mathbf{P}(JohnCalls \mid Burglary = true)$?

$$\mathbf{P}(J \mid b) = \alpha P(b) \sum_{E} P(E) \sum_{A} P(A \mid b, E)\, \mathbf{P}(J \mid A) \sum_{M} P(M \mid A)$$



$$\sum_{M} P(M \mid A) = 1$$

$$\Downarrow$$

The variable M is irrelevant to the query.

Inference with Elimination

- We can remove a leaf node (e.g., M) that is neither a query variable nor an evidence variable.

- After the removal, there may be more leaf nodes that are irrelevant. Remove them as well, and so on.

Inference with Elimination

- We can remove a leaf node (e.g., M) that is neither a query variable nor an evidence variable.

- After the removal, there may be more leaf nodes that are irrelevant. Remove them as well, and so on.

Every variable that is not an ancestor of a query variable or evidence variable is irrelevant to the query.

Inference with Elimination

- We can remove a leaf node (e.g., M) that is neither a query variable nor an evidence variable.

- After the removal, there may be more leaf nodes that are irrelevant. Remove them as well, and so on.

Every variable that is not an ancestor of a query variable or evidence variable is irrelevant to the query.

- Using *reverse topological order* for variables, exact inference with elimination can be 1,000 times faster than the enumeration algorithm.

Inference with Elimination

- We can remove a leaf node (e.g., M) that is neither a query variable nor an evidence variable.

- After the removal, there may be more leaf nodes that are irrelevant. Remove them as well, and so on.

Every variable that is not an ancestor of a query variable or evidence variable is irrelevant to the query.

- ♦ Using *reverse topological order* for variables, exact inference with elimination can be 1,000 times faster than the enumeration algorithm.

- ♦ If we want to compute posterior probabilities for all the variables rather than answer individual queries, we can use clustering algorithms (i.e., join tree algorithms).