

430.457

Introduction to Intelligent Systems

Prof. Songhwai Oh
ECE, SNU

EXACT INFERENCE IN BAYESIAN NETWORKS

Wumpus World (Ch. 7)

Performance measure:

- +1000: climb out of the cave with the gold
- -1000: fall into a pit or eaten by the wumpus
- -1: each action taken
- -10: using up the arrow

Environment:

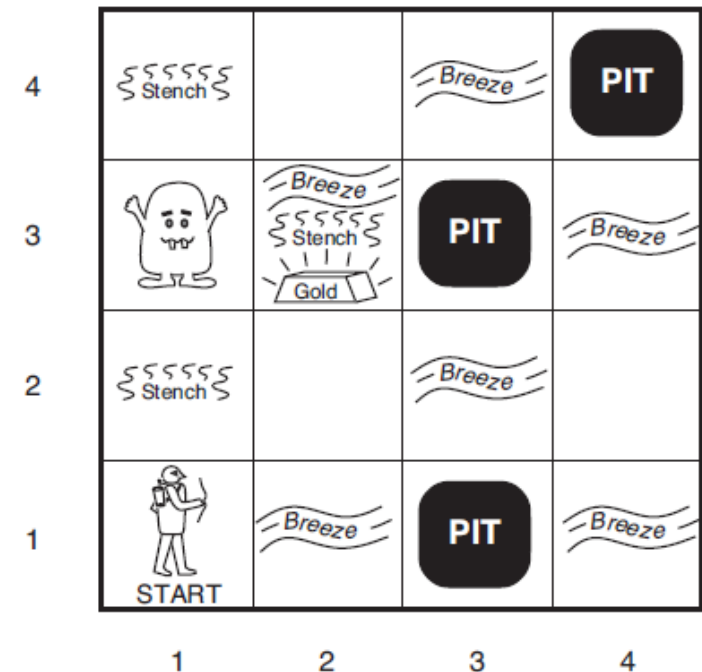
- Locations of gold and wumpus are chosen randomly
- A square has a pit with probability 0.2

Actuators:

- Moves: *Forward*, *Turn Left*, *Turn Right*
- Agent dies if it enters a square with a pit or a live wumpus
- *Grab*: to pick up the gold
- *Shoot*: to fire an arrow (continues until it hits a wall), only one arrow is available
- *Climb*: to climb out of the cave

Sensors:

- *Stench*: in a square adjacent from the wumpus
- *Breeze*: in a square adjacent from a pit
- *Glitter*: in the square where the gold is
- *Bump*: when an agent walks into a wall
- *Scream*: when the wumpus is killed



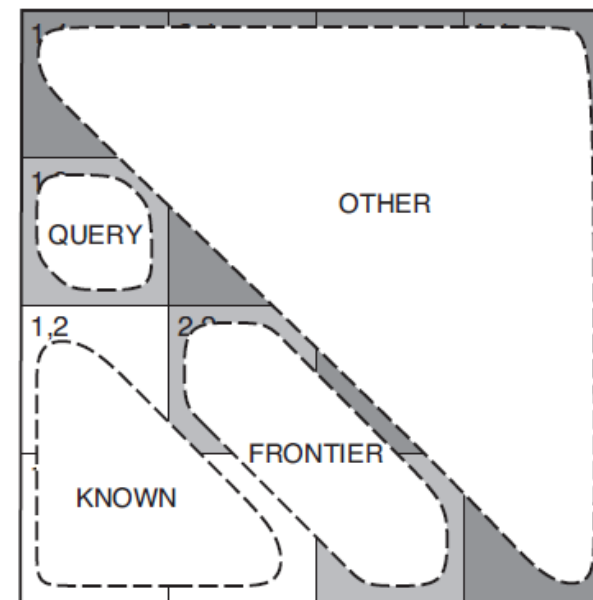
Wumpus World: Probabilistic Inference

For this example, we ignore the wumpus and the gold.

- $P_{i,j}$: True (or 1) if square $[i, j]$ contains a pit
- $B_{i,j}$: True if square $[i, j]$ is breezy
- Known facts and observations: $known = \neg p_{1,1} \wedge \neg p_{1,2} \wedge \neg p_{2,1}, b = \neg b_{1,1} \wedge b_{1,2} \wedge b_{2,1}$
- What is $P(P_{1,3}|known, b)$?

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 B OK	2,2	3,2	4,2
1,1 OK	2,1 B OK	3,1	4,1

(a)

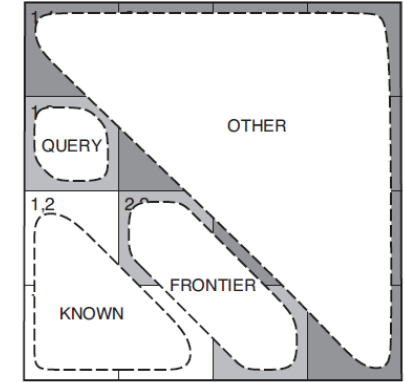


(b)

What is $P(P_{1,3}|known, b)$?

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 B OK	2,2	3,2	4,2
1,1 OK	2,1 B OK	3,1	4,1

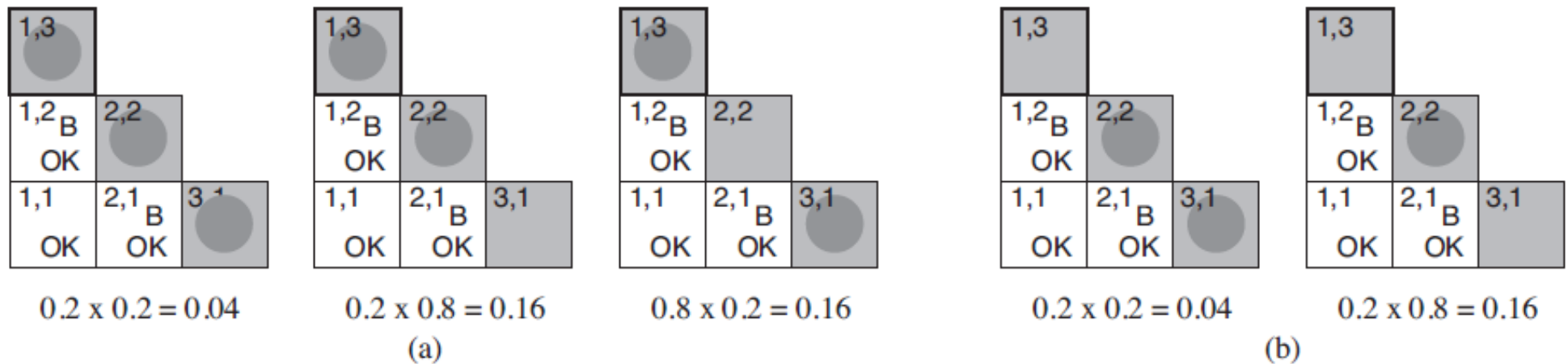
(a)



(b)

$$\begin{aligned}
 P(P_{1,3}|known, b) &= \alpha \sum_{unknown} P(P_{1,3}, known, b, unknown) \\
 &= \alpha \sum_{unknown} P(b|P_{1,3}, known, unknown) P(P_{1,3}, known, unknown) \\
 &= \alpha \sum_{frontier} \sum_{other} P(b|P_{1,3}, known, frontier, other) P(P_{1,3}, known, frontier, other) \\
 &= \alpha \sum_{frontier} \sum_{other} P(b|P_{1,3}, known, frontier) P(P_{1,3}, known, frontier, other) \\
 &= \alpha \sum_{frontier} P(b|P_{1,3}, known, frontier) \sum_{other} P(P_{1,3}, known, frontier, other) \\
 &= \alpha \sum_{frontier} P(b|P_{1,3}, known, frontier) \sum_{other} P(P_{1,3}) P(known) P(frontier) P(other) \\
 &= \alpha P(P_{1,3}) P(known) \sum_{frontier} P(b|P_{1,3}, known, frontier) P(frontier) \sum_{other} P(other) \\
 &= \alpha' P(P_{1,3}) \sum_{frontier} P(b|P_{1,3}, known, frontier) P(frontier)
 \end{aligned}$$

$$P(P_{1,3}|known, b) = \alpha' P(P_{1,3}) \sum_{frontier} P(b|P_{1,3}, known, frontier) P(frontier)$$



$$P(P_{1,3}|known, b) = \alpha' \langle 0.2(0.04 + 0.16 + 0.16), 0.8(0.04 + 0.16) \rangle$$

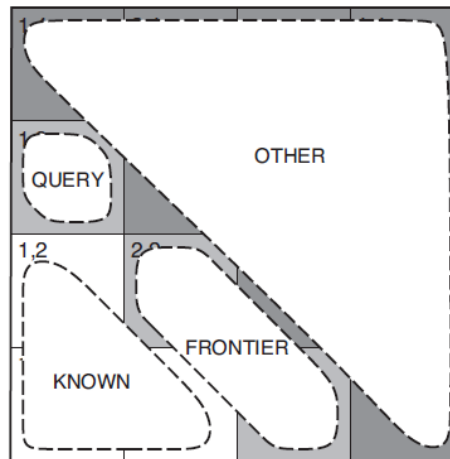
$$\approx \langle 0.31, 0.69 \rangle$$

Probability Inference

- Partition variables in a Bayesian network as $\mathbf{X} = \{X\} \cup \mathbf{E} \cup \mathbf{Y}$
- X , query variables
- $\mathbf{E} = \{E_1, \dots, E_m\}$, evidence variables, and $\mathbf{e}$ is a particular observed event
- $\mathbf{Y} = \{Y_1, \dots, Y_l\}$, hidden variables
- Goal: $P(X|\mathbf{e})$, posterior distribution

1,4	2,4	3,4	4,4
1,3	2,3	3,3	4,3
1,2 B OK	2,2	3,2	4,2
1,1 OK	2,1 B OK	3,1	4,1

(a)

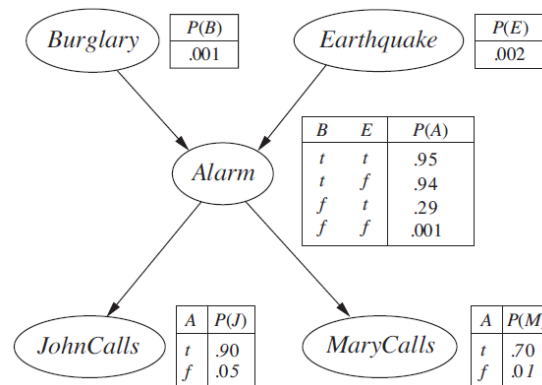


(b)

What is $P(P_{1,3} | \text{known}, b)$?

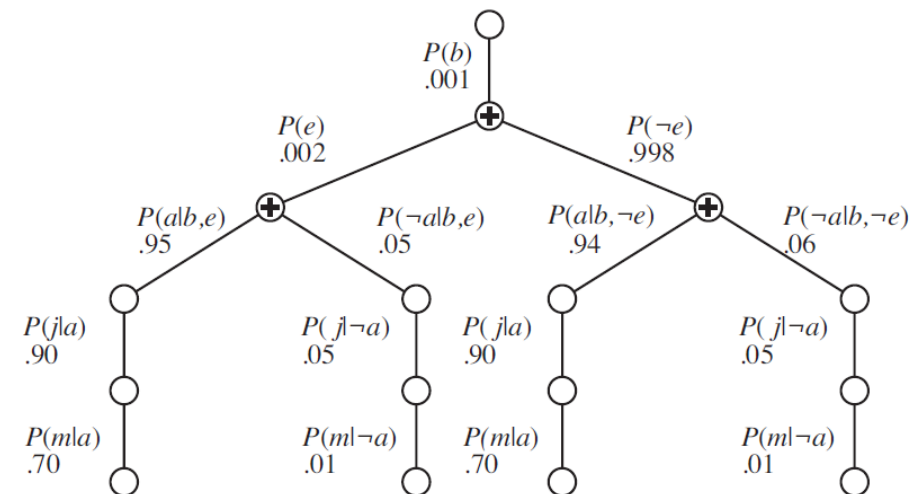
Inference by Enumeration

- $P(X|\mathbf{e}) = \alpha P(X, \mathbf{e}) = \alpha \sum_{\mathbf{y}} P(X, \mathbf{e}, \mathbf{y})$.
- A query can be answered using a Bayesian network by computing sums of products of conditional probabilities from the network.



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

$$\begin{aligned}
 P(b|j, m) &= \alpha P(b, j, m) = \alpha \sum_e \sum_a P(b, j, m, e, a) \\
 &= \alpha \sum_e \sum_a P(b)P(e)P(a|b, e)P(j|a)P(m|a) \\
 &= \alpha P(b) \sum_e P(e) \sum_a P(a|b, e)P(j|a)P(m|a)
 \end{aligned}$$



```

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  $\{X\} \cup \mathbf{E} \cup \mathbf{Y}$   /*  $\mathbf{Y} = \text{hidden variables}$  */

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

```

```

function ENUMERATE-ALL( $vars, \mathbf{e}$ ) returns a real number
  if EMPTY?( $vars$ ) then return 1.0
   $Y \leftarrow$  FIRST( $vars$ )
  if  $Y$  has value  $y$  in  $\mathbf{e}$ 
    then return  $P(y \mid \text{parents}(Y)) \times$  ENUMERATE-ALL(REST( $vars$ ),  $\mathbf{e}$ )
    else return  $\sum_y P(y \mid \text{parents}(Y)) \times$  ENUMERATE-ALL(REST( $vars$ ),  $\mathbf{e}_y$ )
    where  $\mathbf{e}_y$  is  $\mathbf{e}$  extended with  $Y = y$ 

```

For n Boolean variables, the time complexity is $O(2^n)$.

Variable Elimination Algorithm

$$\mathbf{P}(B | j, m) = \alpha \underbrace{\mathbf{P}(B)}_{f_1(B)} \sum_e \underbrace{P(e)}_{f_2(E)} \sum_a \underbrace{\mathbf{P}(a | B, e)}_{f_3(A, B, E)} \underbrace{P(j | a)}_{f_4(A)} \underbrace{P(m | a)}_{f_5(A)}$$

$$f_4(A) = \begin{pmatrix} P(j | a) \\ P(j | \neg a) \end{pmatrix} = \begin{pmatrix} 0.90 \\ 0.05 \end{pmatrix}$$

$$f_5(A) = \begin{pmatrix} P(m | a) \\ P(m | \neg a) \end{pmatrix} = \begin{pmatrix} 0.70 \\ 0.01 \end{pmatrix}$$

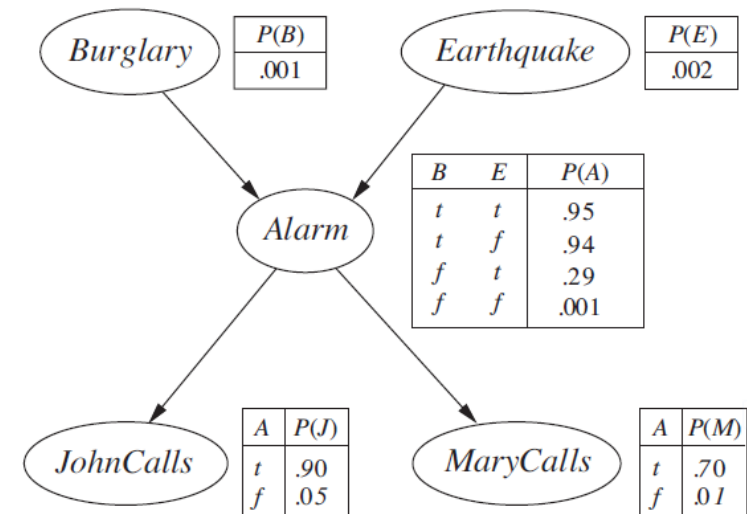
$$\mathbf{P}(B | j, m) = \alpha f_1(B) \times \sum_e f_2(E) \times \sum_a f_3(A, B, E) \times f_4(A) \times f_5(A)$$

$$\begin{aligned} f_6(B, E) &= \sum_a f_3(A, B, E) \times f_4(A) \times f_5(A) \\ &= (f_3(a, B, E) \times f_4(a) \times f_5(a)) + (f_3(\neg a, B, E) \times f_4(\neg a) \times f_5(\neg a)) \end{aligned}$$

$$\mathbf{P}(B | j, m) = \alpha f_1(B) \times \sum_e f_2(E) \times f_6(B, E)$$

$$\begin{aligned} f_7(B) &= \sum_e f_2(E) \times f_6(B, E) \\ &= f_2(e) \times f_6(B, e) + f_2(\neg e) \times f_6(B, \neg e) \end{aligned}$$

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



point-wise product

```

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( $factors$ ))

```

Example:

$$\mathbf{P}(B | j, m) = \alpha \mathbf{f}_1(B) \times \sum_e \mathbf{f}_2(E) \times \sum_a \mathbf{f}_3(A, B, E) \times \mathbf{f}_4(A) \times \mathbf{f}_5(A)$$

$$\mathbf{P}(B | j, m) = \alpha \mathbf{f}_1(B) \times \sum_e \mathbf{f}_2(E) \times \mathbf{f}_6(B, E)$$

$$\mathbf{P}(B | j, m) = \alpha \mathbf{f}_1(B) \times \mathbf{f}_7(B)$$

- The size of a factor is determined by the ordering of variables.
- Determining the optimal ordering is NP-hard