AIFA: Bayesian Network Inference

18/03/2024

Koustav Rudra

Inference using Belief Networks

- Diagnostic inference
  - From effects to causes
  - Given that JohnCalls, infer P(Burglary|JohnCalls)

- Causal inference
  - From causes to effects
  - Given Burglary, infer P(JohnCalls|Burglary)

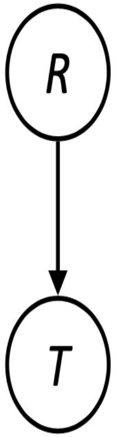
Inference using Belief Networks

- Intercausal inferences (between causes of common effect)
  - Given Alarm, we have P(Burglary|Alarm) = 0.376
  - If we add evidence that earthquake is True then P(Burglary | Alarm ∧ Earthquake) = 0.003

- Mixed inferences
  - Setting the effect JohnCalls to True and cause Earthquake to False
  - P(Alarm|JohnCalls ∧ ~Earthquake) = 0.003

Example: Traffic

- Causal direction



$P(R)$

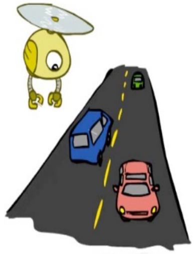
+r	1/4
-r	3/4

$P(T|R)$

+r	+t	3/4
	-t	1/4
-r	+t	1/2
	-t	1/2

$P(T,R)$

+r	+t	3/16
+r	-t	1/16
-r	+t	6/16
-r	-t	6/16

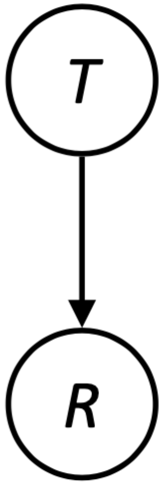


P(R)	
+r	$^1/_4$

R	P(T)
+r	$^3/_4$
-r	$^1/_2$

Example: Traffic

- Reverse causality?



$P(T)$

+t	9/16
-t	7/16

$P(R|T)$

+t	+r	1/3
	-r	2/3
-t	+r	1/7
	-r	6/7

$P(T, R)$

+r	+t	3/16
+r	-t	1/16
-r	+t	6/16
-r	-t	6/16

Probabilistic Inference

- Probabilistic inference: compute a desired probability from other known probabilities (e.g. conditional from joint)

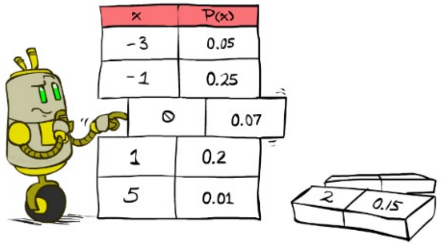
- We generally compute conditional probabilities
  - P(on time | no reported accidents) = 0.90
  - These represent the agent's beliefs given the evidence

- Probabilities change with new evidence:
  - P(on time | no accidents, 5 a.m.) = 0.95
  - P(on time | no accidents, 5 a.m., raining) = 0.80
  - Observing new evidence causes beliefs to be updated

Inference by Enumeration

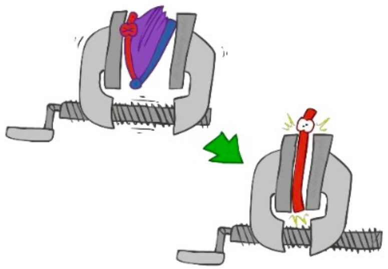
- General Case:
  - Evidence variables: $E_1, E_2, \dots E_k = e_1, e_2, \dots, e_k$
  - Query Variable: Q
  - Hidden variables: $H_1, H_2, \dots H_r$
  - $X_1, X_2, \dots X_n$ all variables
- We want: $P(Q|e_1, e_2, \dots, e_k)$

Step 1: Select the entries consistent with the evidence



$$P(Q, e_1, e_2, \dots, e_k) = \sum_{h_1, h_2, \dots, h_r} P(Q, h_1, h_2, \dots, h_r, e_1, e_2, \dots, e_k)$$

Step 2: Sum out H to get joint of Query and evidence



Step 3: Normalize

$$\times \frac{1}{Z}$$

$$Z = \sum_q P(Q, e_1, e_2, \dots, e_k)$$

$$P(Q|e_1, e_2, \dots, e_k) = \frac{1}{Z} P(Q, e_1, e_2, \dots, e_k)$$

The Chain Rule

- More generally, can always write any joint distribution as an incremental product of conditional distributions
  - $P(x_1, x_2, \ldots, x_n) = \prod_i P(x_i | x_1, x_2, \ldots, x_{i-1})$

Bayes Rule

- Two ways to factor a joint distribution over two variables:
  - $P(x, y) = P(x|y) \times P(y) = P(y|x) \times P(x)$

- Dividing we get
  - $P(x|y) = \frac{P(y|x)}{P(y)} P(x)$

- Why is this useful?
  - Lets us build one conditional from its reverse
  - Often one conditional is tricky but the other one is simple
  - Foundation of many systems (e.g. ASR, MT)

Bayes Rule

R	P
Sun	0.8
Rain	0.2

P(W)

D	W	P
Wet	Sun	0.1
Dry	Sun	0.9
Wet	Rain	0.7
Dry	Rain	0.3

P(D|W)

P(W|Dry)

Inference by Enumeration

Inference using Belief Networks

$$P(B|J) = \frac{P(BJ)}{P(J)}$$

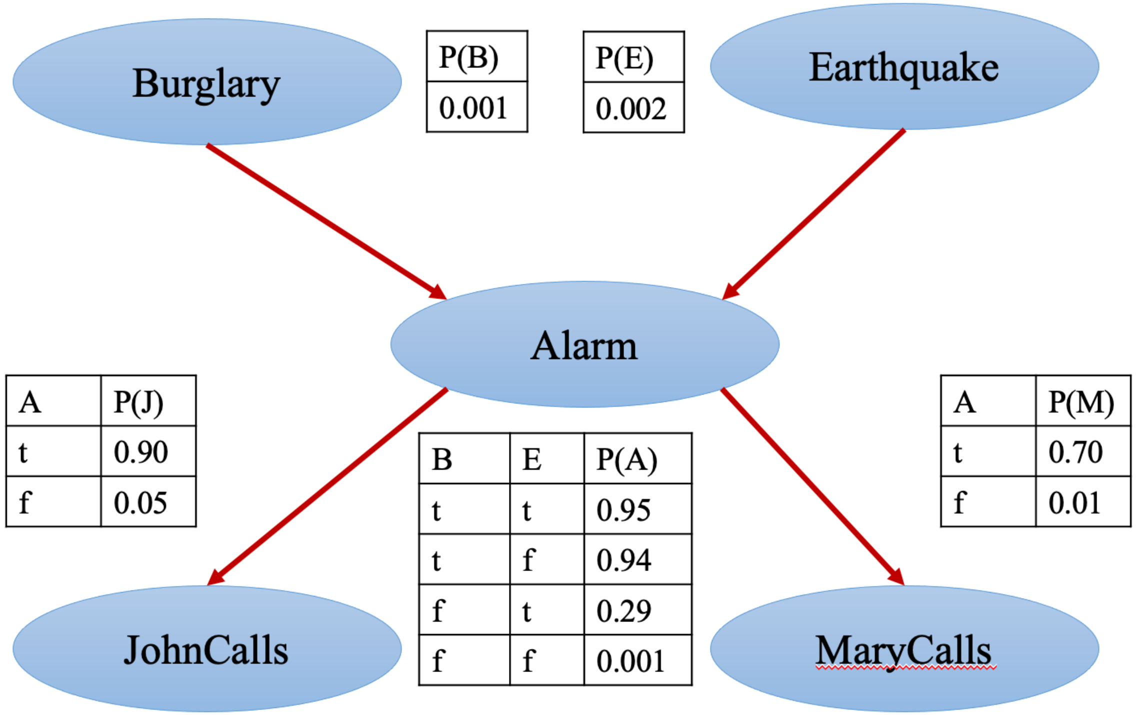
$$P(BJ) = P(BJA) + P(BJA')$$

$$P(BJA) = P(J|AB)P(AB) + P(J|A'B)P(A'B)$$

$$P(AB) = P(ABE) + P(ABE')$$
$$P(AB) = P(A|BE)P(BE) + P(A|BE')P(BE')$$
$$P(AB) = 0.00095$$

$$P(A'B) = 0.00005$$

$$P(BJ) = 0.90 * 0.00095 + 0.05 * 0.00005$$



P(B)
0.001

P(E)
0.002

B	E	P(A)
t	t	0.95
t	f	0.94
f	t	0.29
f	f	0.001

A	P(J)
t	0.90
f	0.05

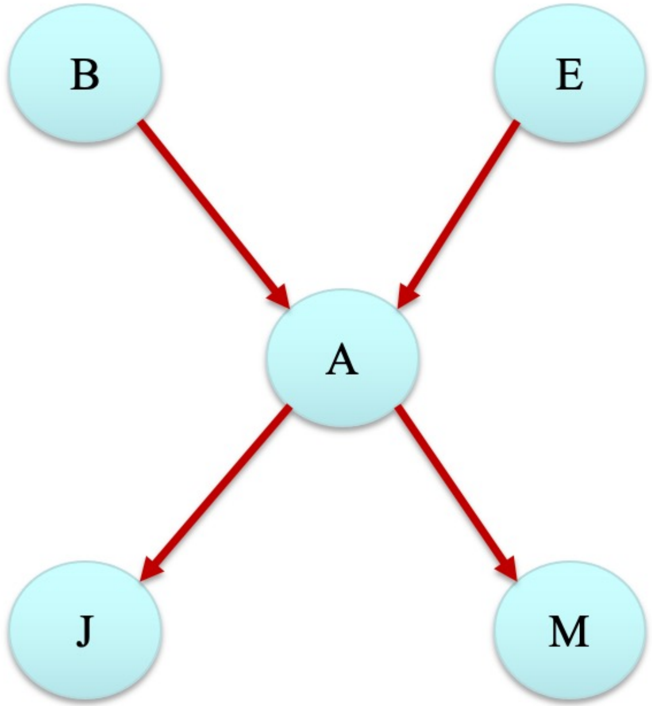
A	P(M)
t	0.70
f	0.01

Bayesian Network: Inference

- Compute posterior probability distribution of a set of query variables
  - Given some observed event [evidence variables]
  - Some unobserved events [Hidden variables]

- P(B|J=True, M=True)

- Hidden variables:
  - Earthquake
  - Alarm



Bayesian Network: Inference by Enumeration

- P(B|J=True, M=True)

- Hidden variables:
  - Earthquake
  - Alarm

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

- Add four terms each computed by multiplying five numbers

- If network contains n variables, complexity O($n2^n$)

Bayesian Network: Inference by Enumeration

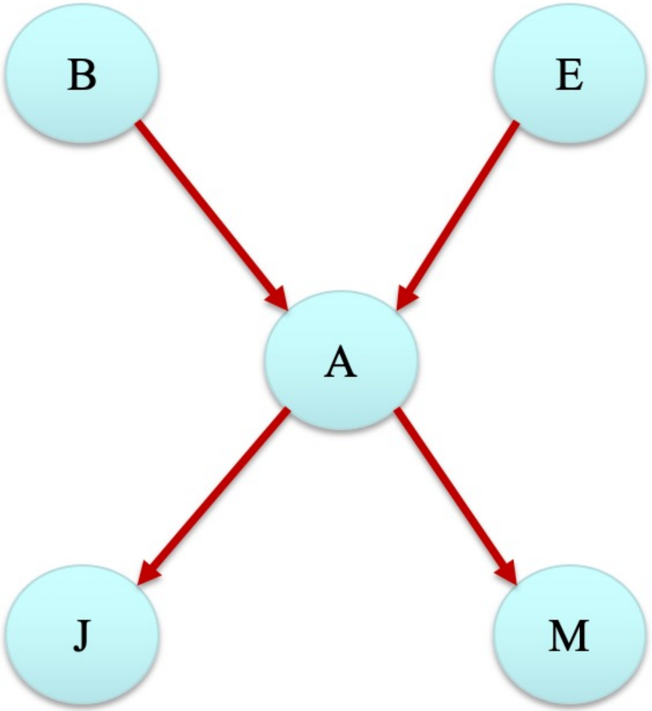
- P(B|J=True, M=True)

- Hidden variables:
  - Earthquake
  - Alarm

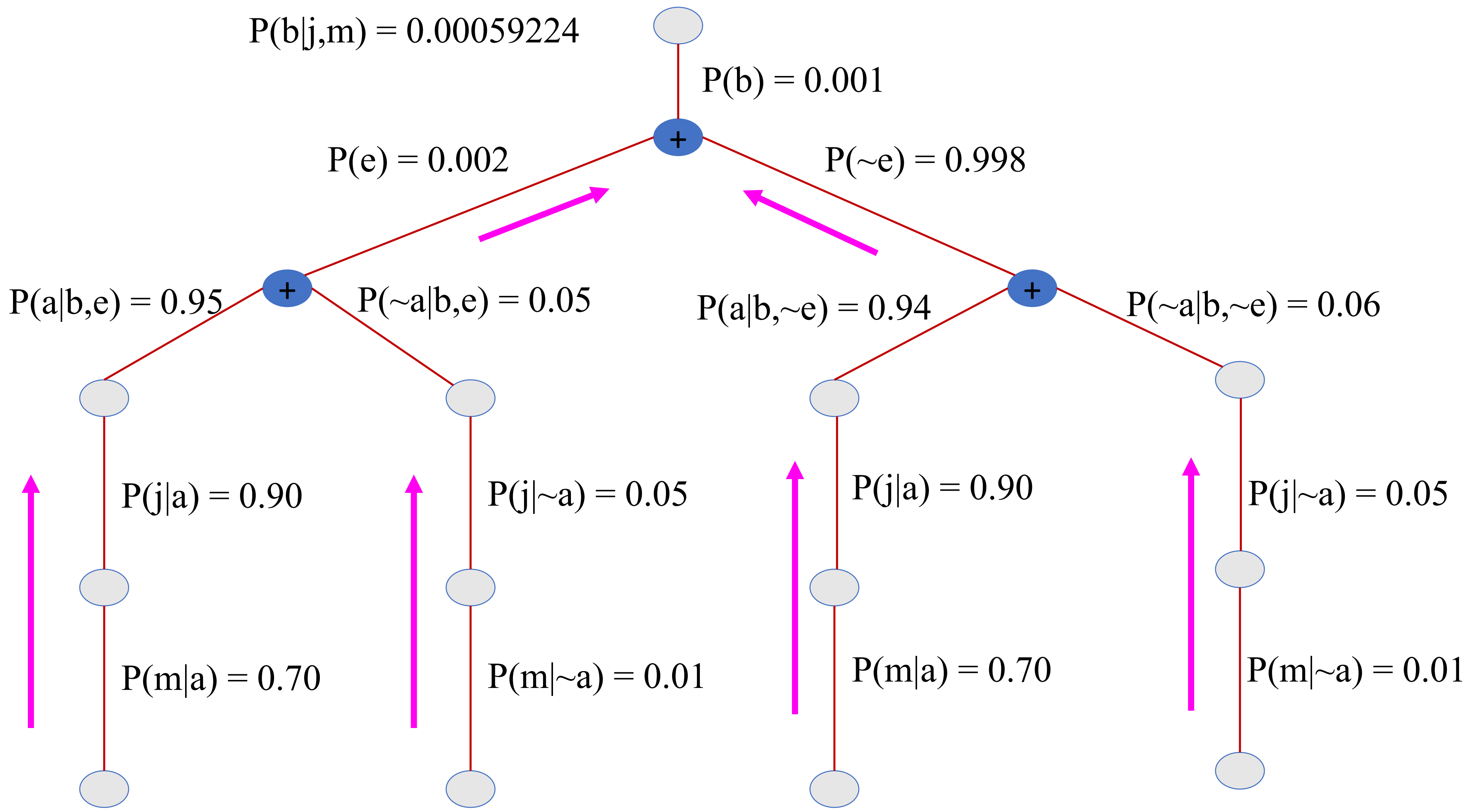
- P(b|j,m) = $\sum_e \sum_a P(b)P(e)P(a|b,e)P(j|a)P(m|a)$

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

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



Bayesian Network: Inference by Enumeration



Bayesian Network: Inference by Enumeration

- P(B|J=True, M=True)

- Hidden variables:
  - Earthquake
  - Alarm

- P(b|j,m) =$P(b)\sum_e P(e)\ \sum_a P(a|b,e)P(j|a)P(m|a)$
- P(b|j,m) = <0.00059224,0.0014919> = <0.284,0.716>

- Chance of a burglary given calls from both neighbours is 28%

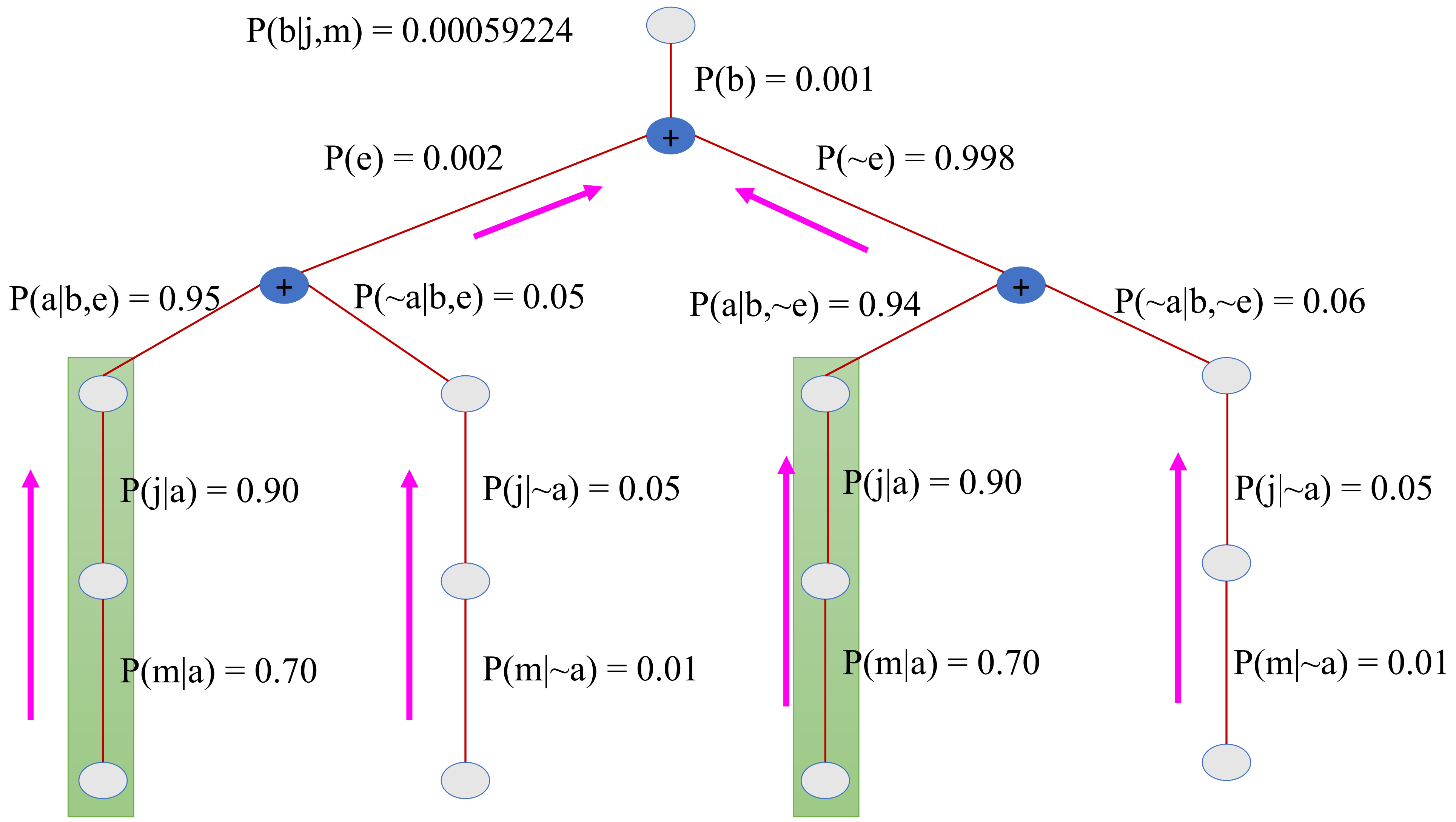
- P(B|J=True, M=True)

- Hidden variables:
  - Earthquake
  - Alarm

- P(b|j,m) =$P(b)\sum_e P(e)\ \sum_a P(a|b,e)P(j|a)P(m|a)$
- P(b|j,m) = <0.00059224,0.0014919> = <0.284,0.716>

- Chance of a burglary given calls from both neighbours is 28%

Bayesian Network: Inference by Enumeration



Thank You