RUTGERS
THE STATE UNIVERSITY
OF NEW JERSEY

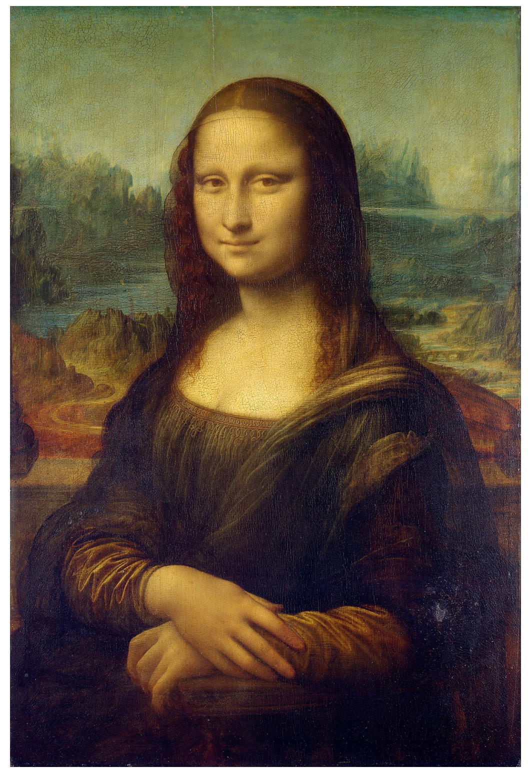
Bayesian Network

Edgar Granados

General Goal of Probabilistic Inference

- **H -** Hypotheses, non directly observable

- **D -** Data



- Compute $P(h \mid d)$

Applying of Bayes

Example

- Marie is getting married tomorrow, at an outdoor ceremony in the desert. In recent years, it has rained only 5 days each year. Unfortunately, the weatherman has predicted rain for tomorrow. When it actually rains, the weatherman correctly forecasts rain 90% of the time. When it doesn't rain, he incorrectly forecasts rain 10% of the time. What is the probability that it will rain on the day of Marie's wedding?

R: Rain or not ; F: Forecast rain or not

$\boxed{\frac{360}{365}}$ $1 - \frac{5}{365}$

$$P(R|F) = \frac{P(F|R)P(R)}{P(F)}$$

$$= \frac{P(F|R)\,P(R)}{P(F|R)P(R) + P(F|\neg R)P(\neg R)} = \frac{.9 \cdot \frac{5}{365}}{.9 \cdot \frac{5}{365} + .1 \cdot \frac{360}{365}}$$

$$= 0.11$$

Applying of Bayes

Example

$1 - 99\%$

- Suppose a drug test is 99% sensitive and 99% specific. That is, the test will produce 99% true positive results for drug users and 99% true negative results for non-drug users. Suppose that 0.5% of people are users of the drug. If a randomly selected individual tests positive, what is the probability he or she is a user?

$$P(\text{User} \mid \text{Positive}) = \frac{P(P \mid U)\, P(U)}{P(P)}$$

$$= \frac{P(P \mid U)\, P(U)}{P(P \mid U)\, P(U) + P(P \mid \neg U)\, P(\neg U)}$$

$$= \frac{.99 \cdot 0.005}{.99 \cdot .005 + 0.01 \cdot .995} = .33$$

General Pipeline for Learning and Inference

MAP & ML

- Pick the best hypothesis given data:

$$h = \arg\max_{h} P(H=h \mid D=d)$$

$$= \arg\max_{h} \frac{P(D=d \mid H=h)\, P(H=h)}{P(D=d)}$$

Max a Posteriori

$$\underbrace{P(H=h \mid D=d)}_{\text{Posterior}} \propto \underbrace{P(D=d \mid H=h)}_{\text{Likelihood}}\, \underbrace{P(H=h)}_{\text{Prior}}$$

Maximum Likelihood (ML)

$$P(H=h \mid D=d) \propto P(D=d \mid H=h)$$

Naïve Bayes

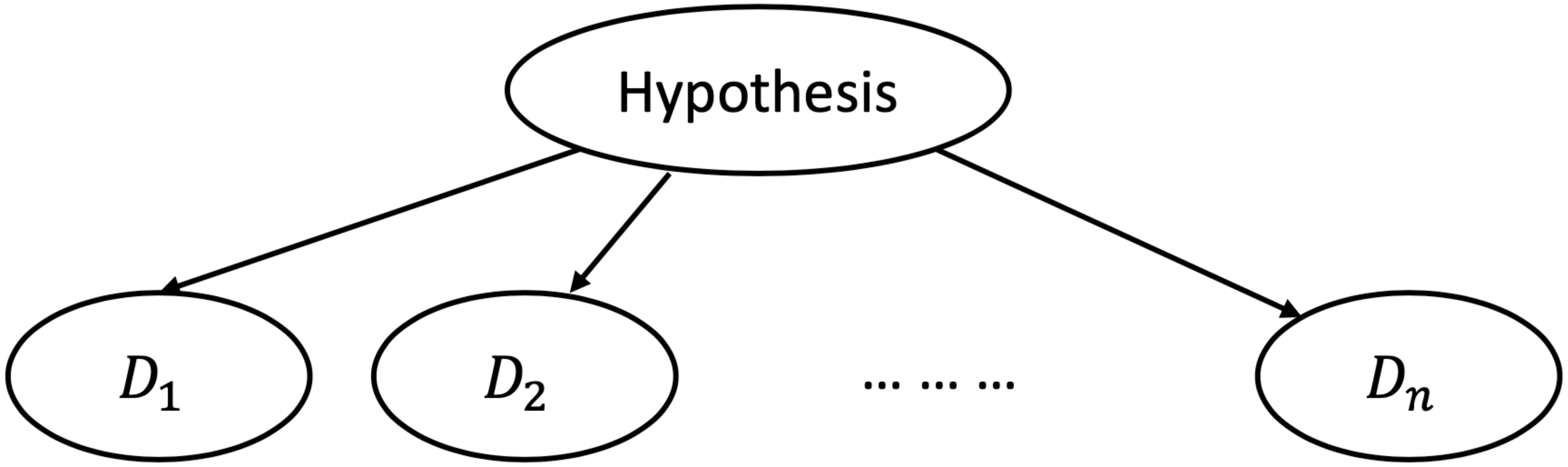
Conditionally Independent given H

- Suppose $\{D_1, (\ldots), D_n\}$ is available
  - Income, age, debt, ...
  - Determine if someone should be approved for a credit

$$P(h \mid d_1, \ldots, d_n) \propto P(d_1, \ldots, d_n)\, P(h) = P(h) \cdot \prod_{1 \le i \le n} P(d_i \mid h)$$

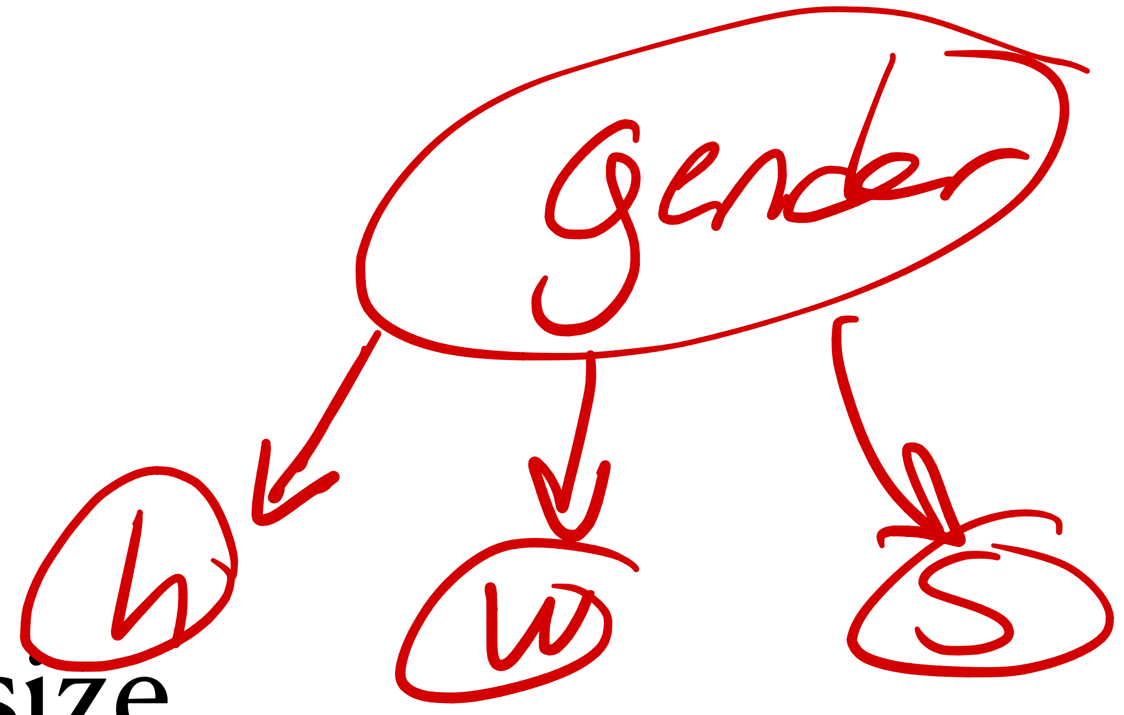
Naïve Bayes

Graphical Model



Naïve Bayes

Example



- Deciding gender based on height, weight, and shoe size

	Height		Weight		Shoe Size	
Gender	μ	σ	μ	σ	μ	σ
M	5.855	0.035	176.25	122.9	11.25	9.2
F	5.4175	0.097	132.5	558.3	7.5	16.7

$N(\mu, \sigma)$

$P(\theta)$?

- Given (h=6,w=130,s=8), what is the predicted gender?

$P(f \mid h=6, w=130, S=8) \propto P(f)\,P(h=6 \mid f)\,P(w=130 \mid f)\,P(s=8 \mid f) = \boxed{5.4 \times 10^{-4}}$

$P(m \mid h=6, w=130, s=8) \propto P(m)\,P(h=6 \mid m)\,P(w=130 \mid m)\,P(s=8 \mid m) = 6.2 \times 10^{-9}$

Naïve Bayes

Example

$W_i \sim$ Word i

- Filter Spam Emails

$$P(class \mid doc) \propto P(class) \prod_i P(W_i \mid class)$$

$$P(spam) = P(\neg spam) = 0.5$$

Email is spam if

$$P(spam \mid document) > P(\neg spam \mid doc)$$

Insanely Easy Way To Become Rich In 2021

What if I told you that you can become a millionaire this year?

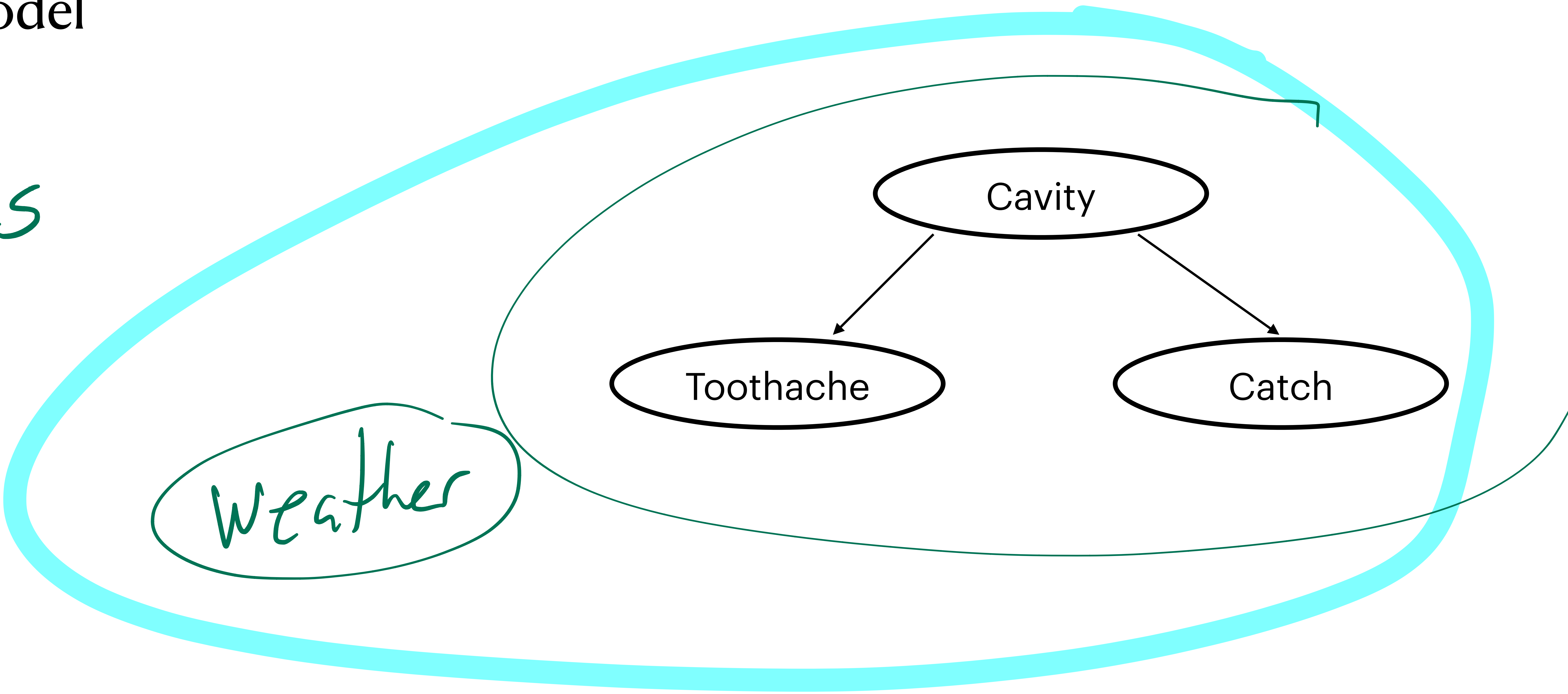
Sounds crazy, right?

You have been chosen to participate in our Loyalty Program for FREE! It will take you only a minute to receive this fantastic prize.

Structure of Bayesian Network

- A type of graphical model
- Nodes – RV
- Directed Edges – Arcs

• Acyclic



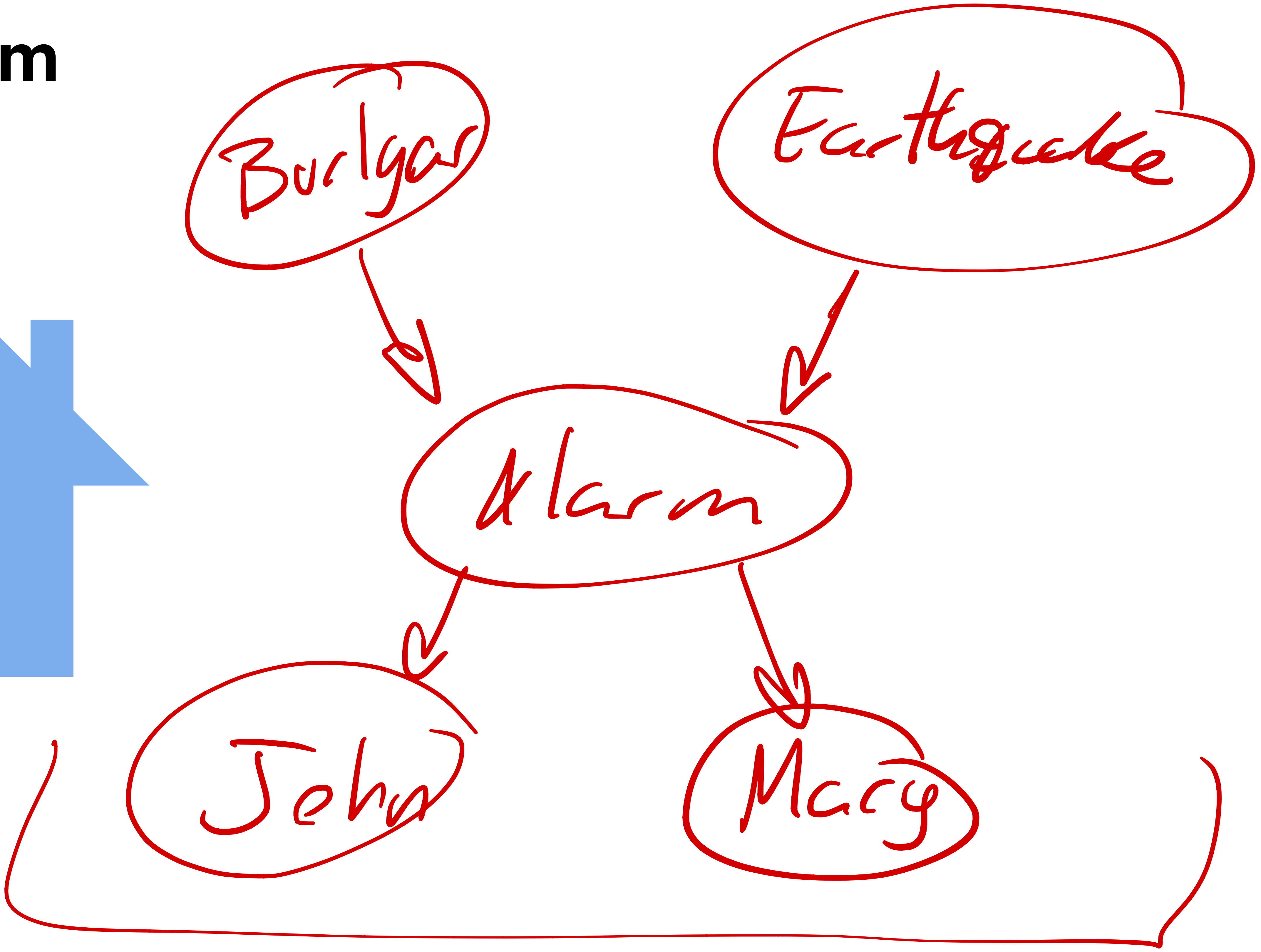
Structure of Bayesian Network

- The arcs determine the structure of a Bayesian Network
  - Tossing n coins?

Bayesian Network

Example: Alarm

Burlgar
Earthquake



Bayesian Network

Conditional Independence

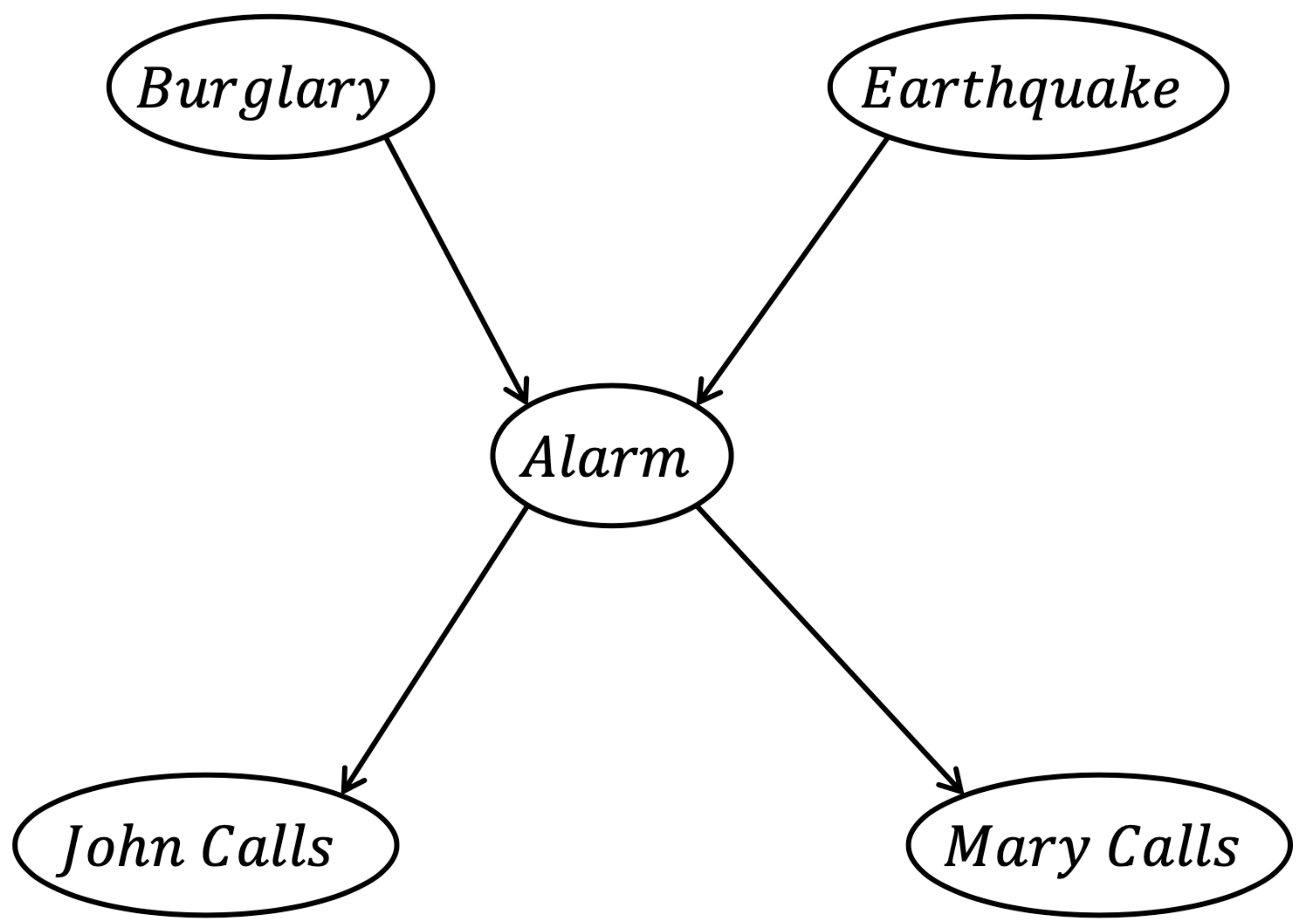
- A node is independent of its non descendants given its parents

$X_1, \dots, X_n$

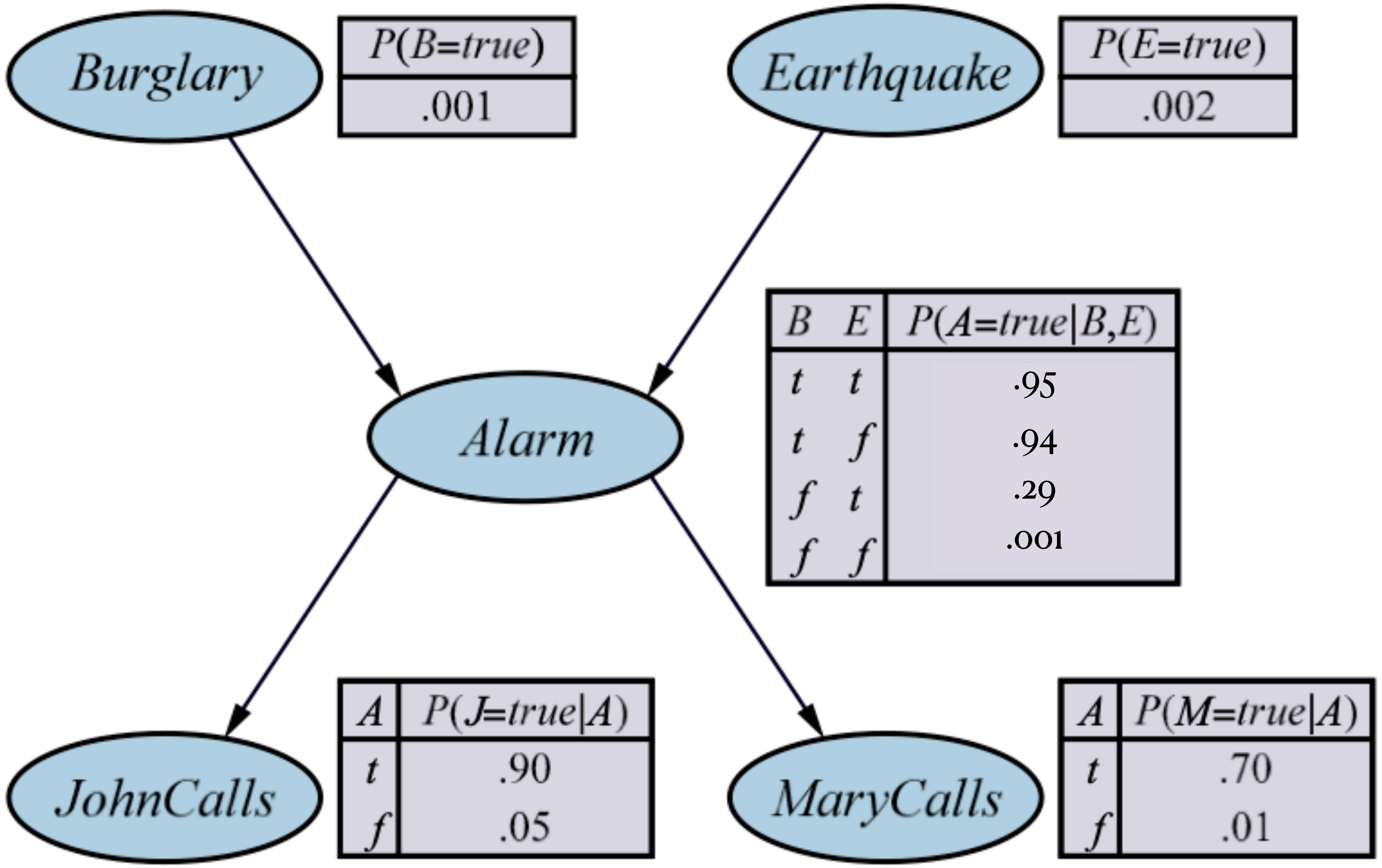
if X_i is parent of $X_j \Rightarrow i < j$

$$P(X_1, \dots, X_n) = \prod_{i=1}^{n} P(X_i \mid X_1, \dots, X_{i-1})$$

$$= \prod_{i=1}^{n} P(X_i \mid \text{Parents}(X_i))$$



Full Joint Distribution



Computing Using CPTs

Example

CPT – conditional Prob. Table

- Compute $P(j, m, a, \neg b, \neg e)$

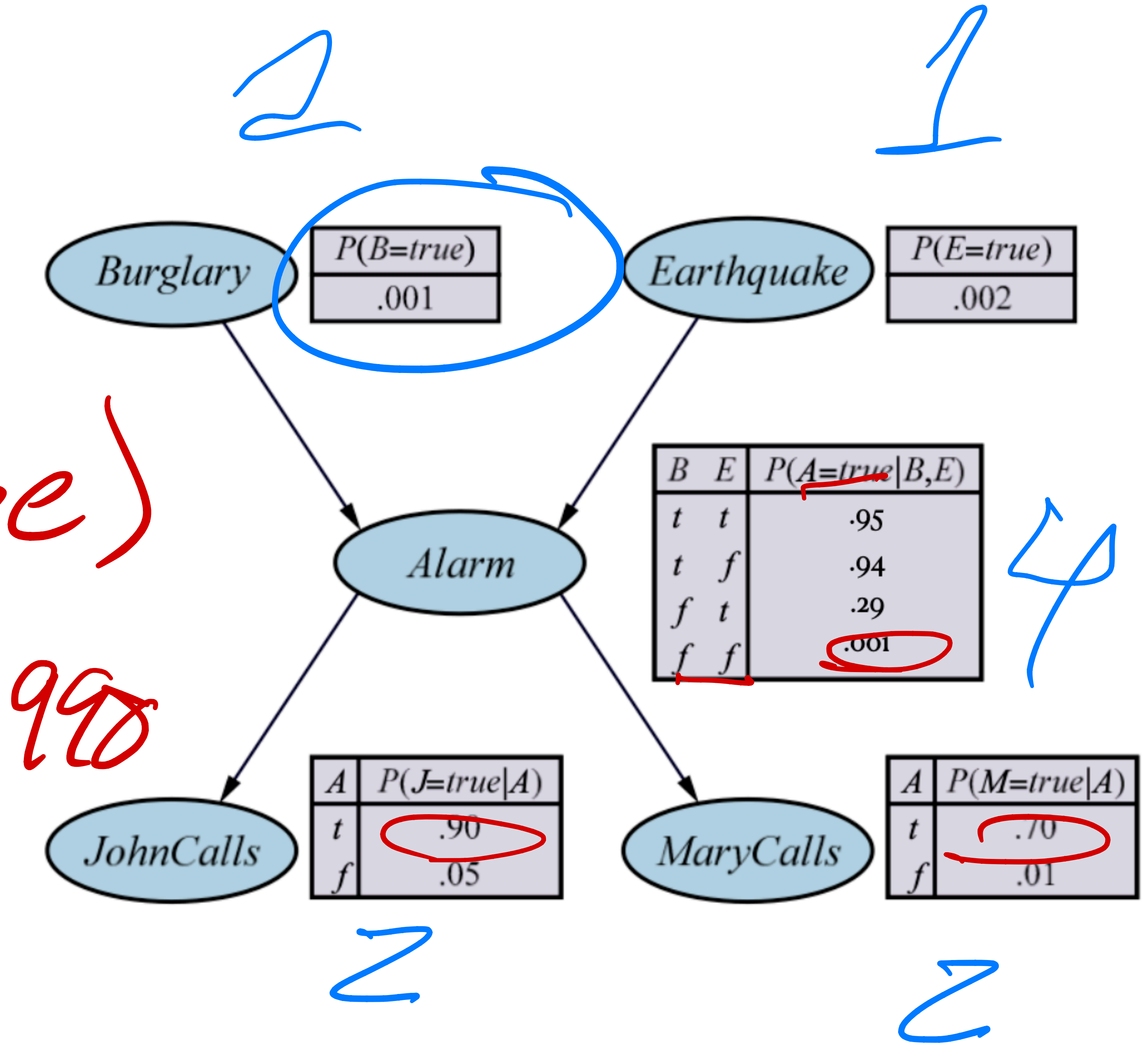
$= P(j,m \mid a,\neg b,\neg e)\,P(a,\neg b,\neg e)$

$= P(j,m \mid a)\,P(a \mid \neg b,\neg e)\,P(\neg b)\,P(\neg e)$

$= P(j \mid a)\,P(m \mid a)\,P(a \mid \neg b,\neg e)\,P(\neg b)\,P(\neg e)$

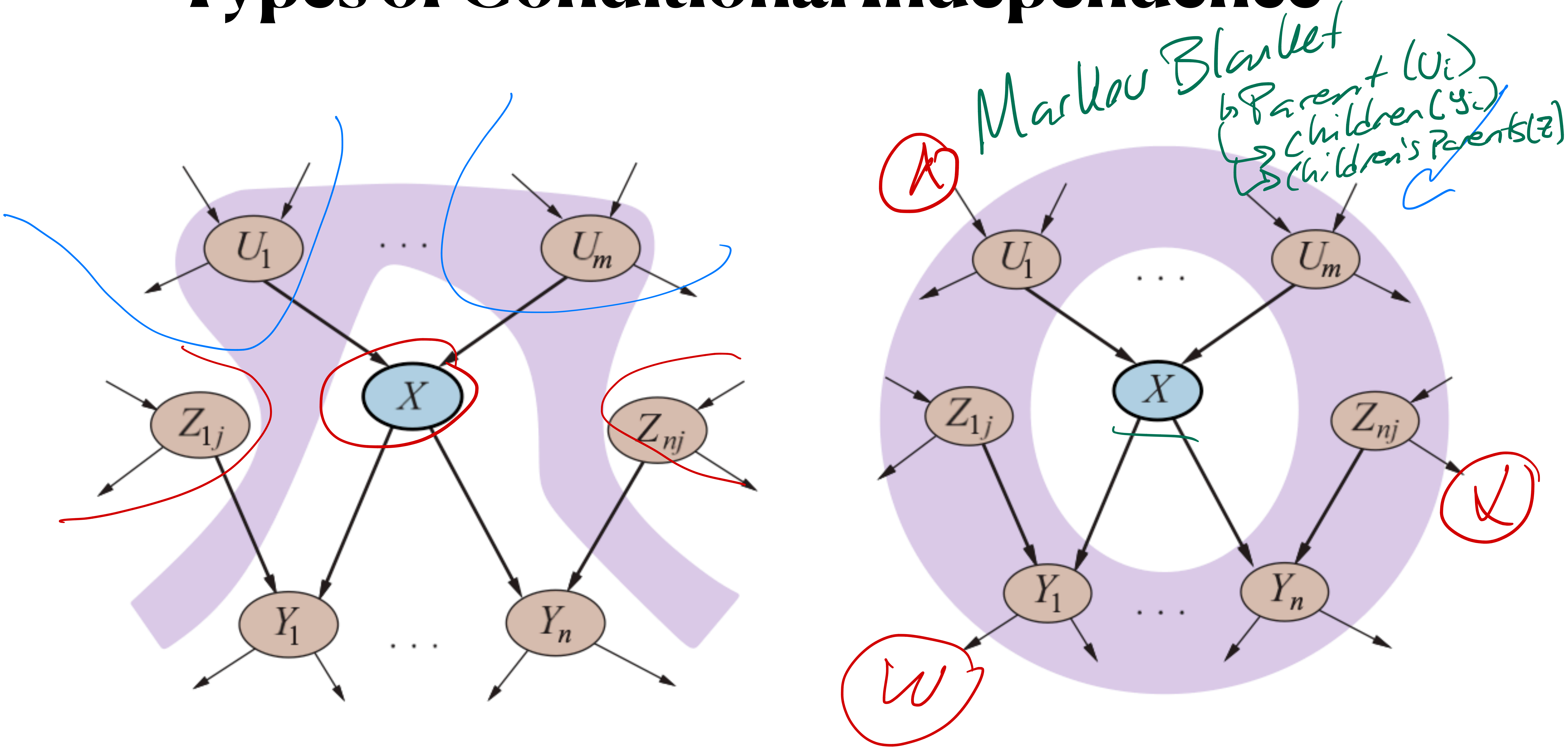
$= .9 \cdot .7 \cdot .001 \cdot .999 \cdot .998$

$\simeq 0.000628$



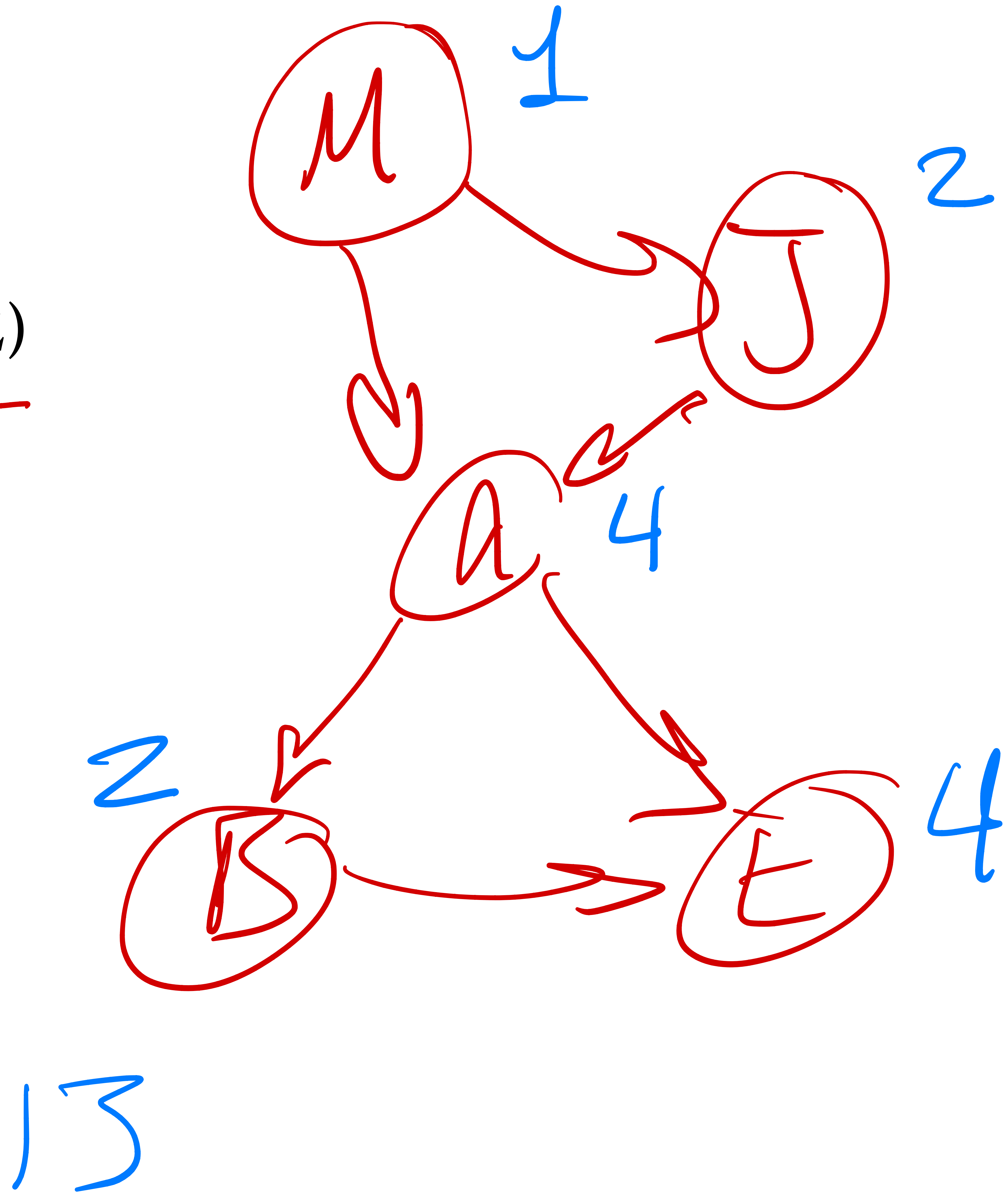
10

Types of Conditional Independence

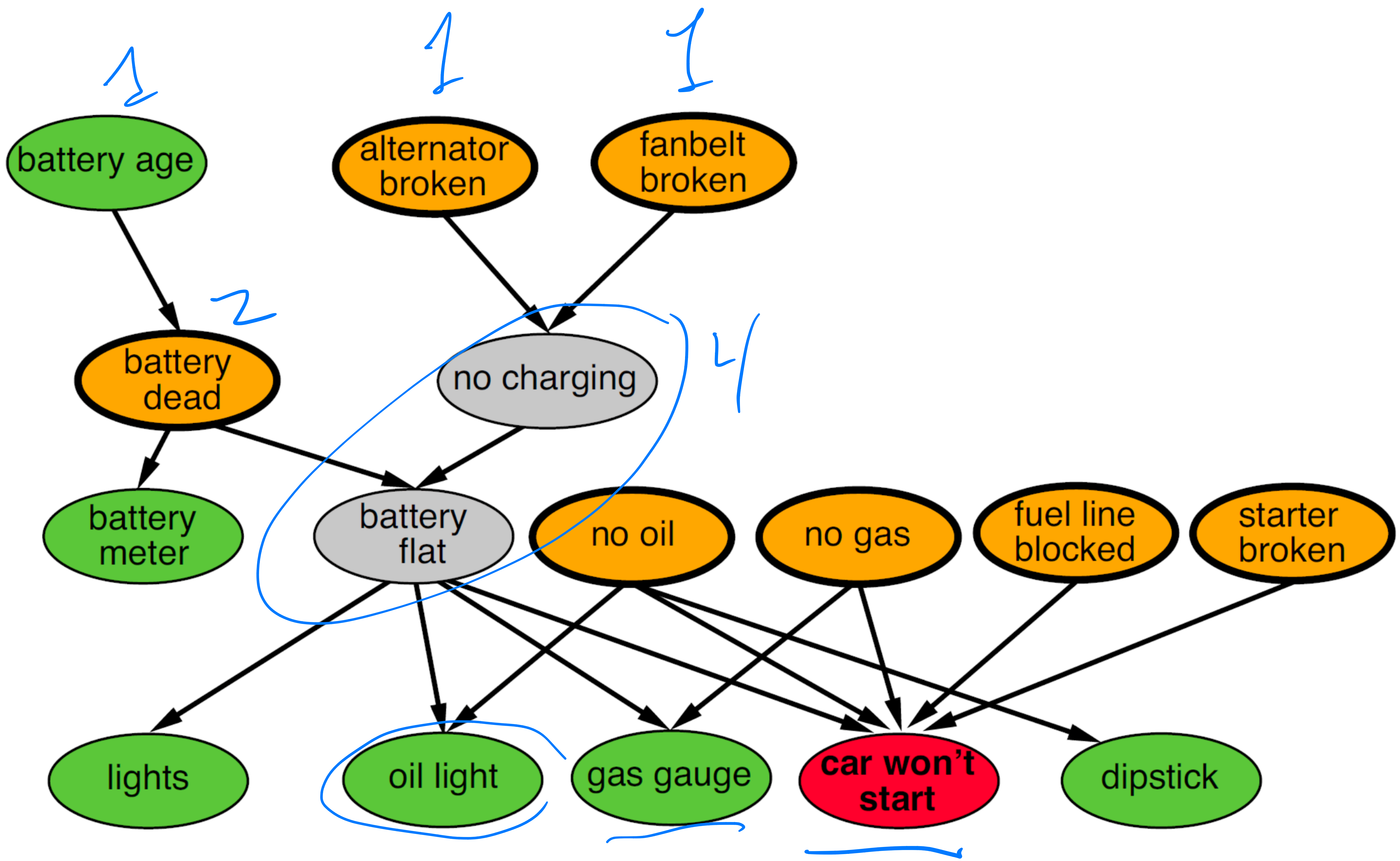


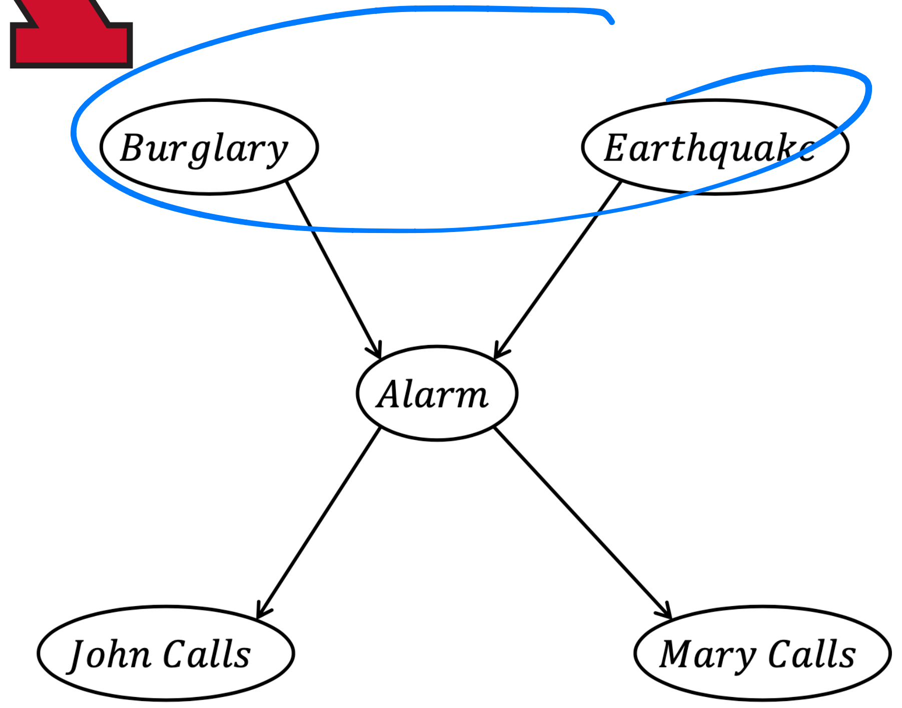
Constructing A Bayesian Network

- ORDER MATTERS
  - What about the sequence (M,J, A, B, E)



Constructing A Bayesian Network





Exact Inference: Enumeration

$$P(X|e) = \frac{P(x,e)}{P(e)} = \alpha P(X,e)$$

- General inference over Bayesian network

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

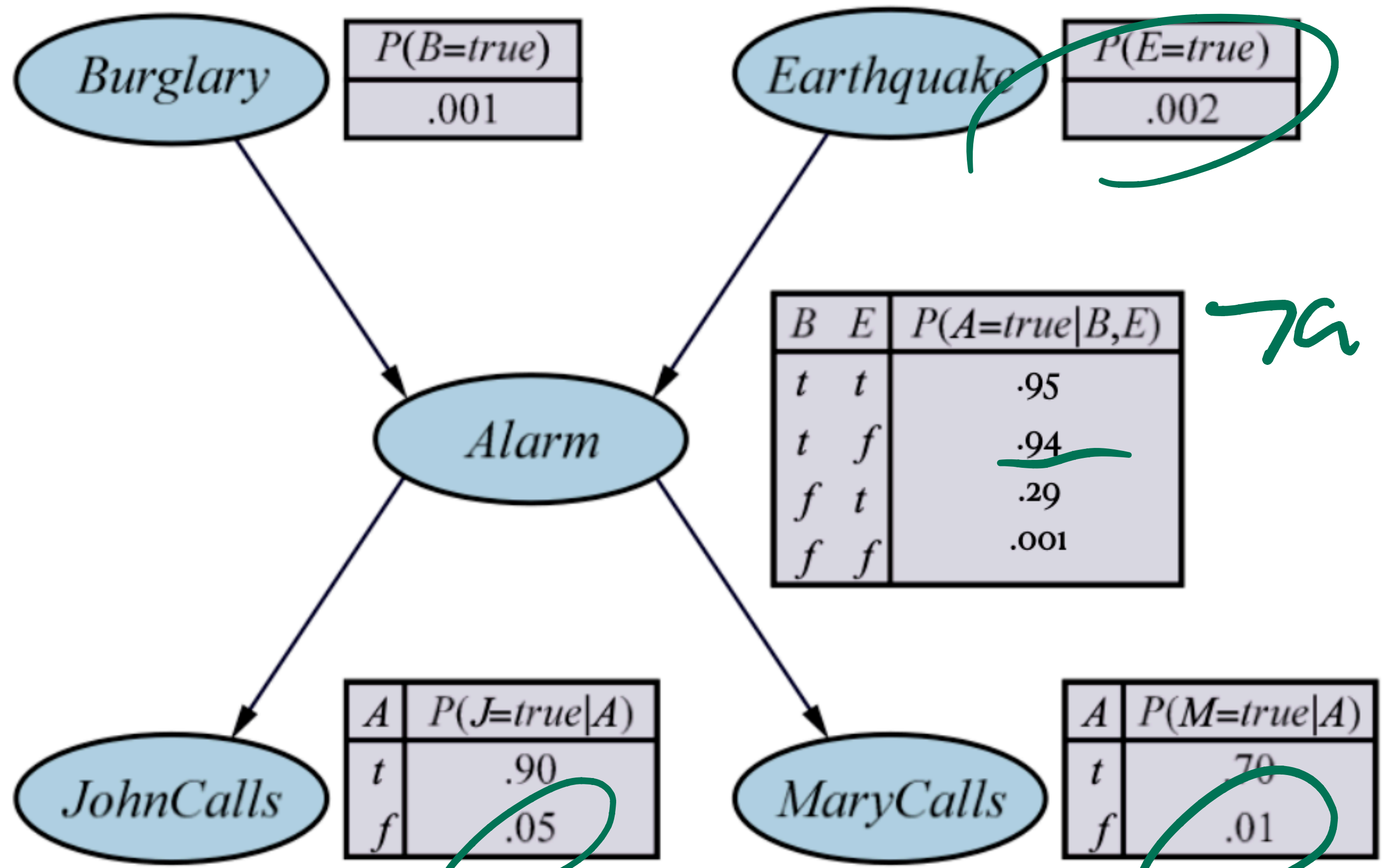
$$P(b|m,j) + P(\neg b|m,j) = 1 = \alpha\left(P(b,m,j) + P(\neg b,m,j)\right)$$

- E.g.:
  - $P(B\,|\,j,m) = \alpha P(B,j,m) = \alpha \sum_e \sum_a P(B,j,m,e,a)$

$B=b$

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

$$P(\neg b|j,m)$$



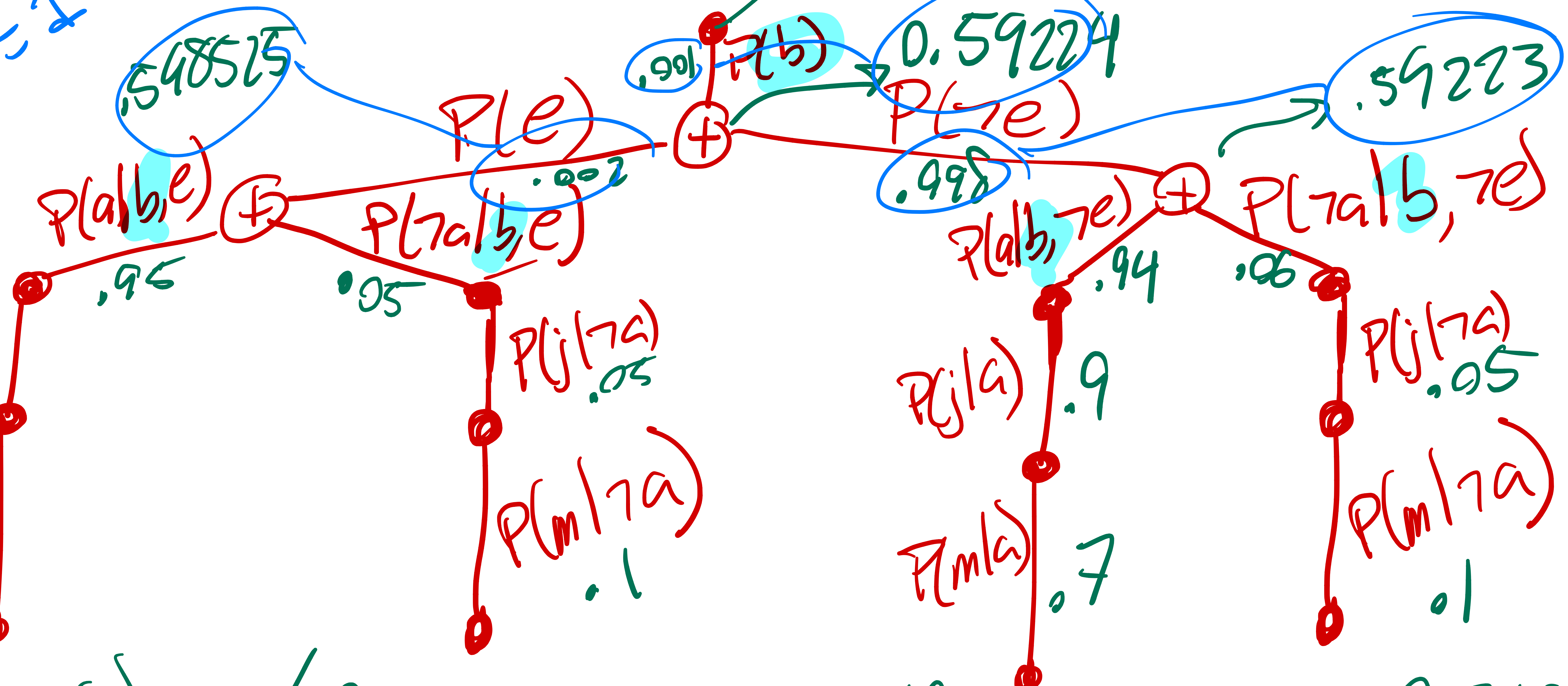
Exact Inference

Enumeration

- DFS traversal through $P(b)\sum_e P(e)\sum_a P(a|b,e)P(j|a)P(m|a)$

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

$P(b|mj) + P(\neg b|mj) = 1$



$P(B|m,j) = \alpha\langle 0.00059224, 0.001491\rangle = \langle 0.284, 0.716\rangle$

Exact Inference

Variable Elimination

- Define **factors** $P(B|j,m) = \alpha \underset{f_1(B)}{P(b)} \sum_{e} \underset{f_2(E)}{P(e)} \sum_{a} \underset{f_3(A,B,E)}{P(a|b,e)} \overset{f_4(A)}{P(j|a)} \overset{f_5(A)}{P(m|a)}$

$$f_1(B) = \begin{pmatrix} P(b) \\ P(\neg b) \end{pmatrix} = \begin{pmatrix} .001 \\ .999 \end{pmatrix} \qquad f_2(E) = \begin{pmatrix} P(e) \\ P(\neg e) \end{pmatrix}$$

$$f_4(A) = \begin{pmatrix} P(j|a) \\ P(j|\neg a) \end{pmatrix} \qquad f_5(A) = \begin{pmatrix} P(m|a) \\ P(m|\neg a) \end{pmatrix}$$

$f_3(A, B, E)$? $\rightarrow$ $2 \times 2 \times 2$

Operations on Factors

- Using **pointwise product** $(\times)$

$$P(B|j,m) = \alpha P(b) \sum_{e} P(e) \sum_{a} P(a|b,e)P(j|a)P(m|a)$$

$$= \alpha f_1(B) \times \sum_{e} f_2(E) \times \sum_{a} f_3(A,B,E) \times f_4(A) \times f_5(A)$$

$f_i \times f_j$ yields a new factor f

$$f(a,b,c) = f_i(a,b)\, f_j(b,c)$$

Pointwise Product

Example

- $f_3(A, B, C) = f_1(A, B) \times f_2(B, C)$

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	$.1 \times .4 = .04$

$$f(B,C) = \sum_a f_3(A,B,C)$$

$$= f_3(a,B,C) + f_3(\neg a,B,C)$$

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

$$= \begin{pmatrix} .24 & .96 \\ .48 & .32 \end{pmatrix}$$

Variable Elimination

- Using **pointwise product**

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

$$= \alpha f_1(B) \times \sum_e f_2(E) \times \sum_a f_3(A,B,E) \times f_4(A) \times f_5(A)$$