CS 2710 Foundations of AI
Lecture 17

Inference in Bayesian belief networks

Milos Hauskrecht
milos@cs.pitt.edu
5329 Sennott Square

Modeling uncertainty with probabilities

- **Knowledge based system era (70s – early 80's)**
  - **Extensional non-probabilistic models**
  - Solve the space, time and acquisition bottlenecks in probability-based models
  - froze the development and advancement of KB systems and contributed to the slow-down of AI in 80s in general
- Breakthrough (late 80s, beginning of 90s)
  - **Bayesian belief networks**
    - Give solutions to the space, acquisition bottlenecks
    - Partial solutions for time complexities
- Bayesian belief network

Bayesian belief networks (BBNs)

Bayesian belief networks.

- Represent the full joint distribution over the variables more compactly with a **smaller number of parameters**.
- Take advantage of **conditional and marginal independences** among random variables

- **A and B are independent**

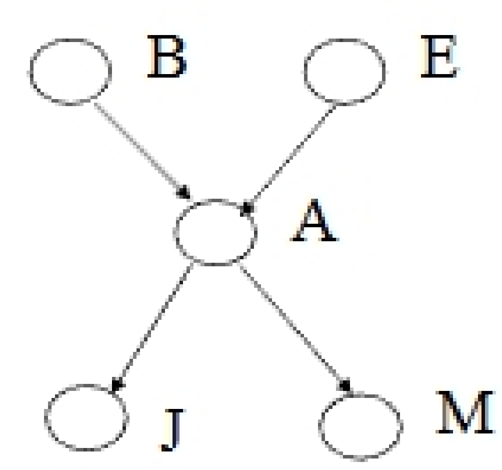
$$P(A,B) = P(A)P(B)$$

- **A and B are conditionally independent given C**

$$P(A,B \mid C) = P(A \mid C)P(B \mid C)$$
$$P(A \mid C,B) = P(A \mid C)$$

Bayesian belief networks (general)

Two components: $B = (S, \Theta_S)$



- **Directed acyclic graph**
  - Nodes correspond to random variables
  - (Missing) links encode independences
- **Parameters**
  - Local conditional probability distributions for every variable-parent configuration

$$\mathbf{P}(X_i \mid pa(X_i))$$

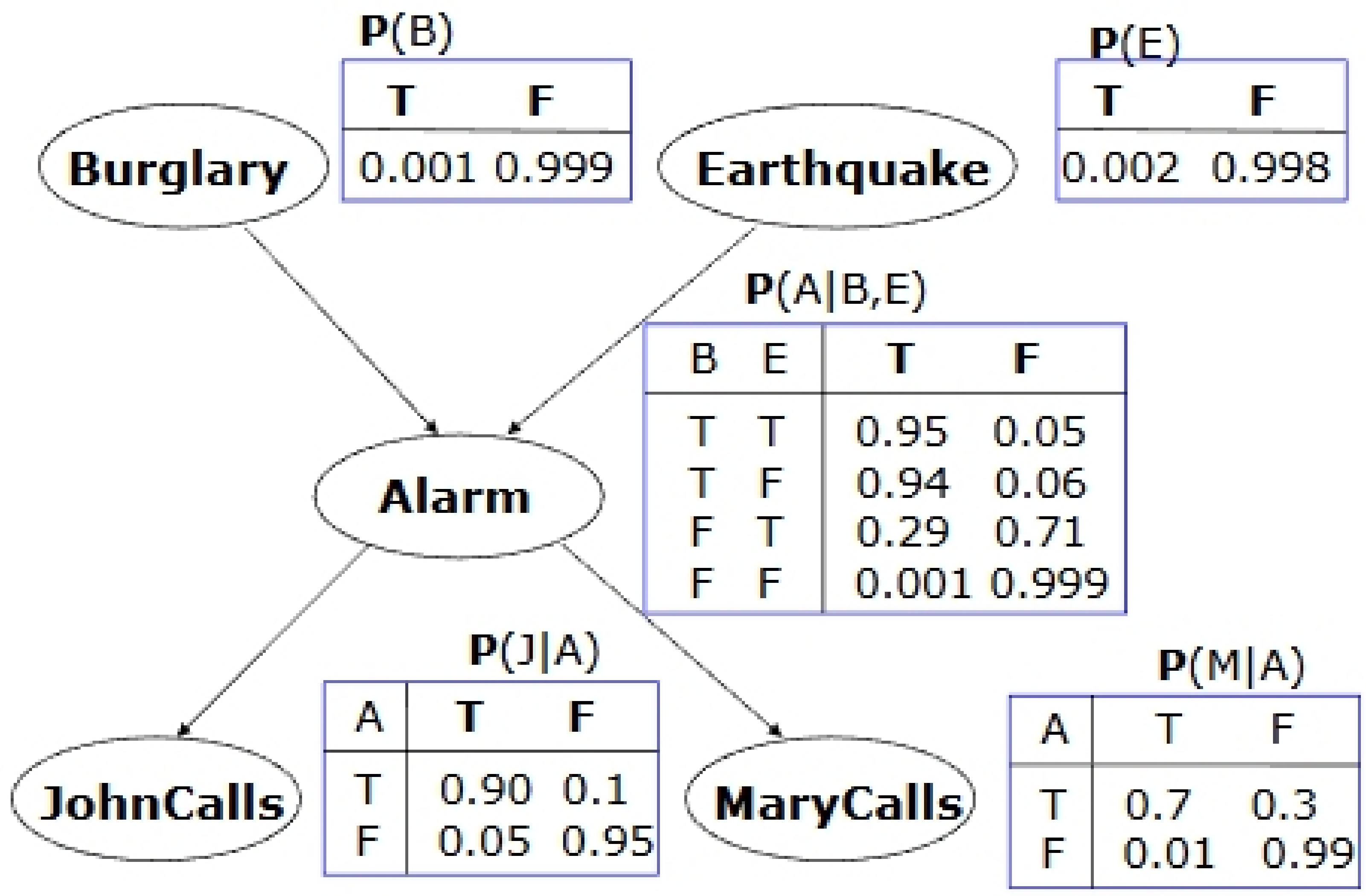
Where:

$pa(X_i)$ - stand for parents of X_i

P(A|B,E)

B	E	**T**	**F**
T	T	0.95	0.05
T	F	0.94	0.06
F	T	0.29	0.71
F	F	0.001	0.999

Bayesian belief network.



P(B)

T	F
0.001	0.999

P(E)

T	F
0.002	0.998

P(A|B,E)

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

P(J|A)

A	T	F
T	0.90	0.1
F	0.05	0.95

P(M|A)

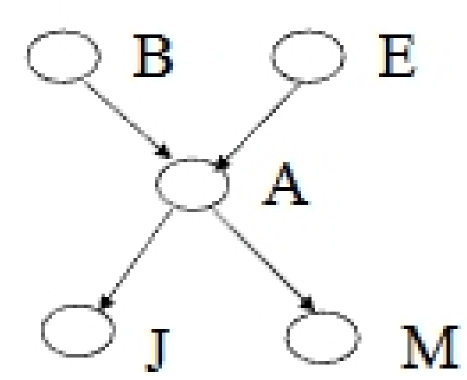
A	T	F
T	0.7	0.3
F	0.01	0.99

Full joint distribution in BBNs

Full joint distribution is defined in terms of local conditional distributions (obtained via the chain rule):

$$\mathbf{P}(X_1, X_2, .., X_n) = \prod_{i=1,..n} \mathbf{P}(X_i \mid pa(X_i))$$

Example:



Assume the following assignment of values to random variables

$B=T, E=T, A=T, J=T, M=F$

Then its probability is:

$P(B=T, E=T, A=T, J=T, M=F) =$

$$P(B=T)P(E=T)P(A=T \mid B=T, E=T)P(J=T \mid A=T)P(M=F \mid A=T)$$