Reasoning under Uncertainty

The intelligent way to handle the unknown

COURSE: CS60045

Pallab Dasgupta
Professor,
Dept. of Computer Sc & Engg

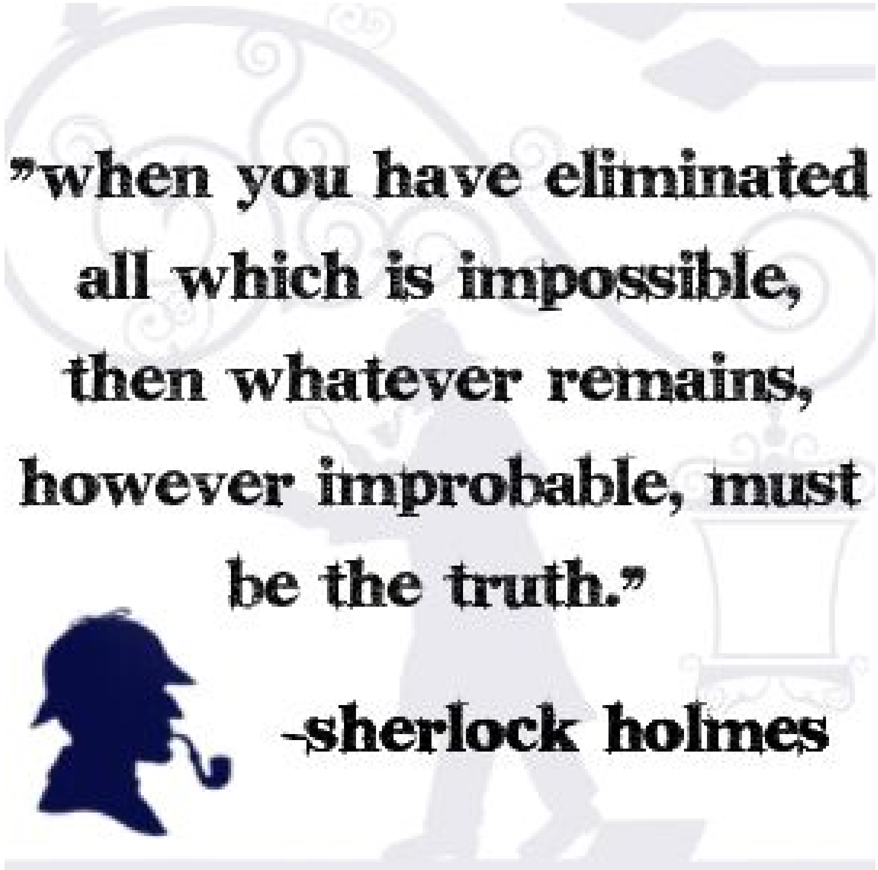
Logical Deduction versus Induction

DEDUCTION

- Commonly associated with *formal logic*
- Involves reasoning from known premises to a conclusion
- The conclusions reached are inevitable, certain, inescapable

INDUCTION

- Commonly known as *informal logic* or *everyday argument*
- Involves drawing uncertain inferences based on probabilistic reasoning
- The conclusions reached are probable, reasonable, plausible, believable



Handling uncertain knowledge

- Classical first order logic has no room for uncertainty

$\forall$p Symptom(p, Toothache) $\Rightarrow$ Disease(p, Cavity)

  - Not correct – toothache can be caused in many other cases
  - In first order logic we have to include all possible causes

    $\forall$p Symptom(p, Toothache) $\Rightarrow$ Disease(p, Cavity) $\vee$ Disease(p, GumDisease) $\vee$ Disease(p, ImpactedWisdom) $\vee$ …

- Similarly, Cavity does not always cause Toothache, so the following is also not true

  $\forall$p Disease(p, Cavity) $\Rightarrow$ Symptom(p, Toothache)

Reasons for using probability

- Specification becomes too large
  - It is too much work to list the complete set of antecedents or consequents needed to ensure an exception-less rule
- Theoretical ignorance
  - The complete set of antecedents is not known
- Practical ignorance
  - The truth of the antecedents is not known, but we still wish to reason

Predicting versus Diagnosing

- Probabilistic reasoning can be used for predicting outcomes (*from cause to effect*)
  - Given that I have a cavity, what is the chance that I will have toothache?

- Probabilistic reasoning can also be used for diagnosis (*from effect to cause*)
  - Given that I am having toothache, what is the chance that it is being caused by a cavity?

We need a methodology for reasoning that can work both ways.

Axioms of Probability

1. All probabilities are between 0 and 1: $0 \leq P(A) \leq 1$
2. $P(True) = 1$ and $P(False) = 0$
3. $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$

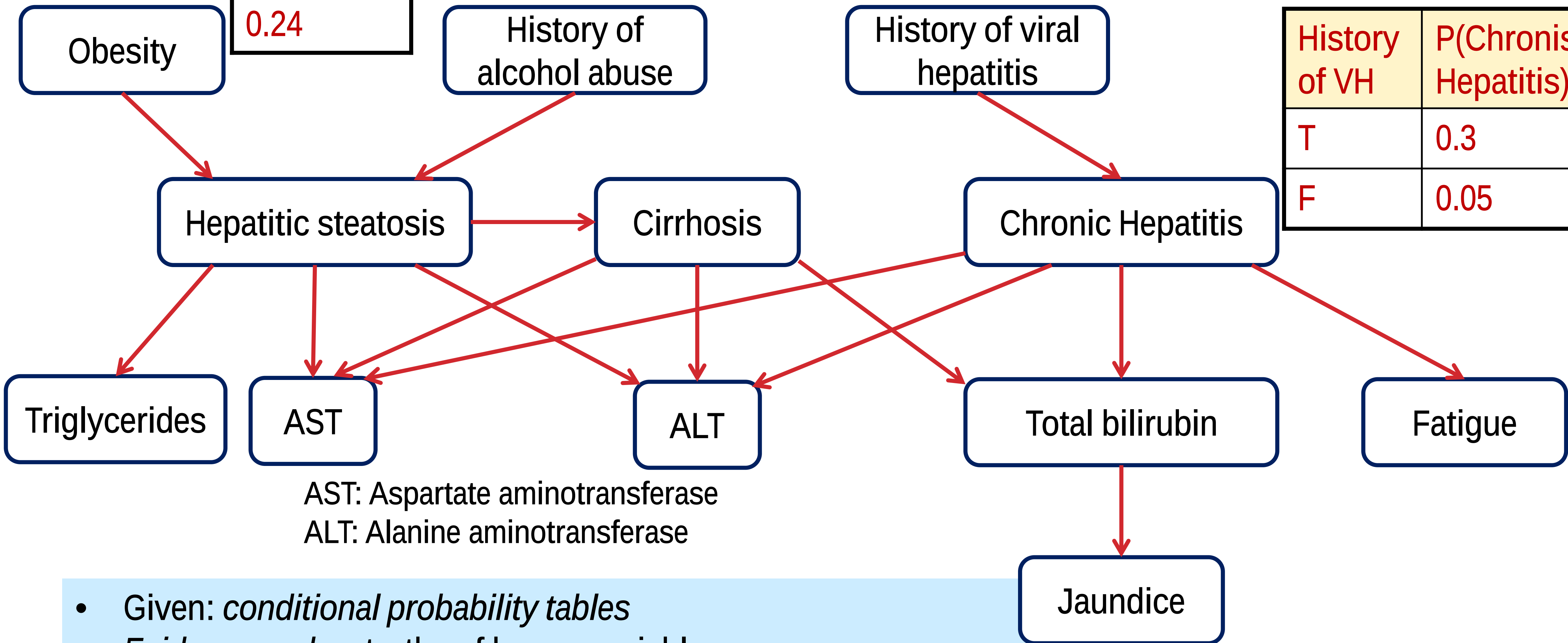
Bayes' Rule

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

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

$$P(B \mid A) = \frac{P(A \mid B)\, P(B)}{P(A)}$$

Bayesian Belief Network



P(Obesity)
0.24

History of VH	P(Chronis Hepatitis)
T	0.3
F	0.05

AST: Aspartate aminotransferase
ALT: Alanine aminotransferase

- Given: *conditional probability tables*
- *Evidence nodes:* truths of known variables
- Goal: *Find probabilities of other variables and/or their combinations*

Belief Networks

A belief network is a graph with the following:

- Nodes: Set of random variables
- Directed links: The intuitive meaning of a link from node X to node Y is that X has a direct influence on Y

Each node has a conditional probability table that quantifies the effects that the parent have on the node.

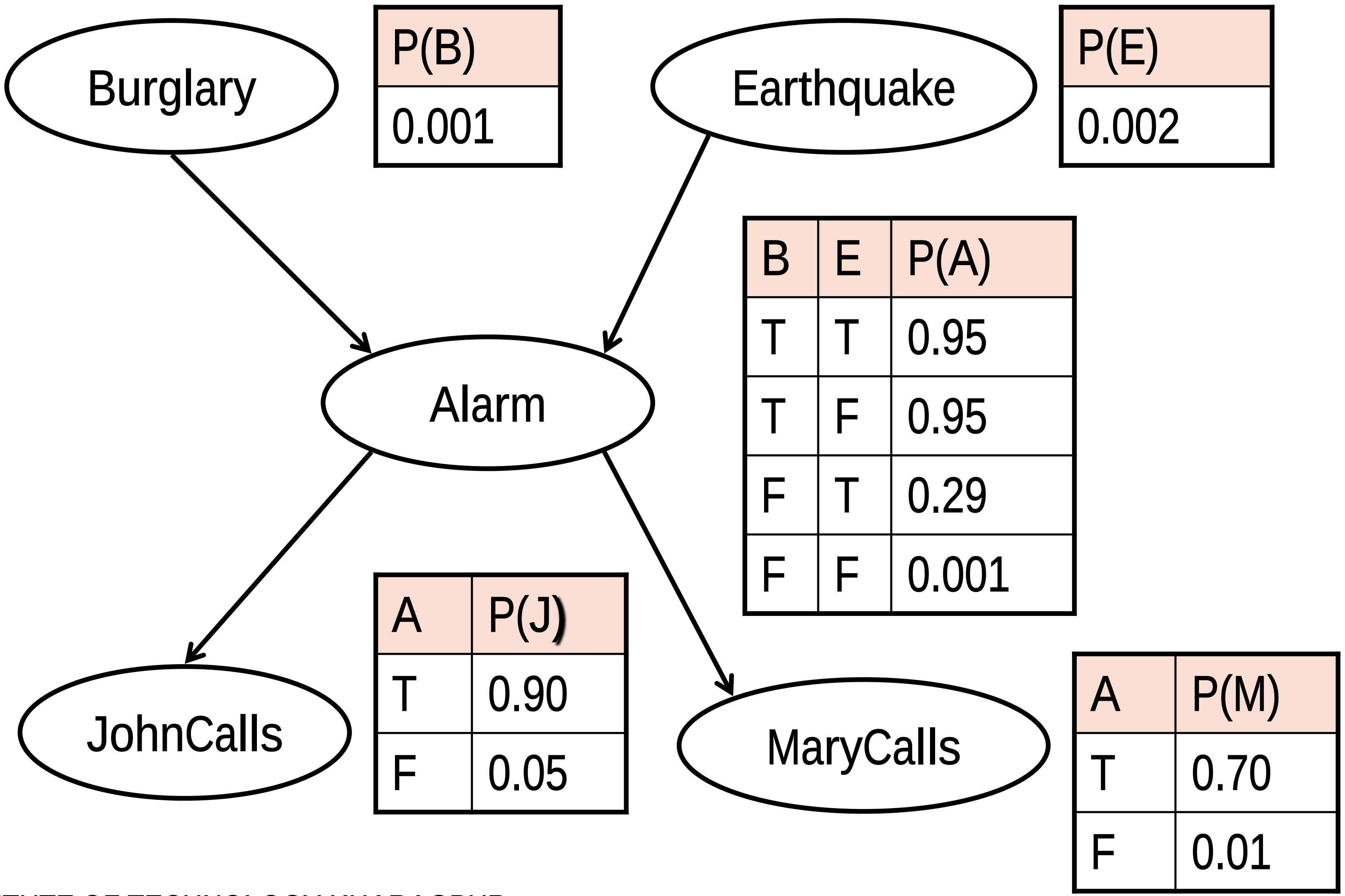
The graph has no directed cycles. It is a *directed acyclic graph* (DAG).

Classical Example

- Burglar alarm at home
  - Fairly reliable at detecting a burglary
  - Responds at times to minor earthquakes

- Two neighbors, on hearing alarm, calls police
  - John always calls when he hears the alarm, but sometimes confuses the telephone ringing with the alarm and calls then, too.
  - Mary likes loud music and sometimes misses the alarm altogether

Belief Network Example



P(B)
0.001

P(E)
0.002

B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

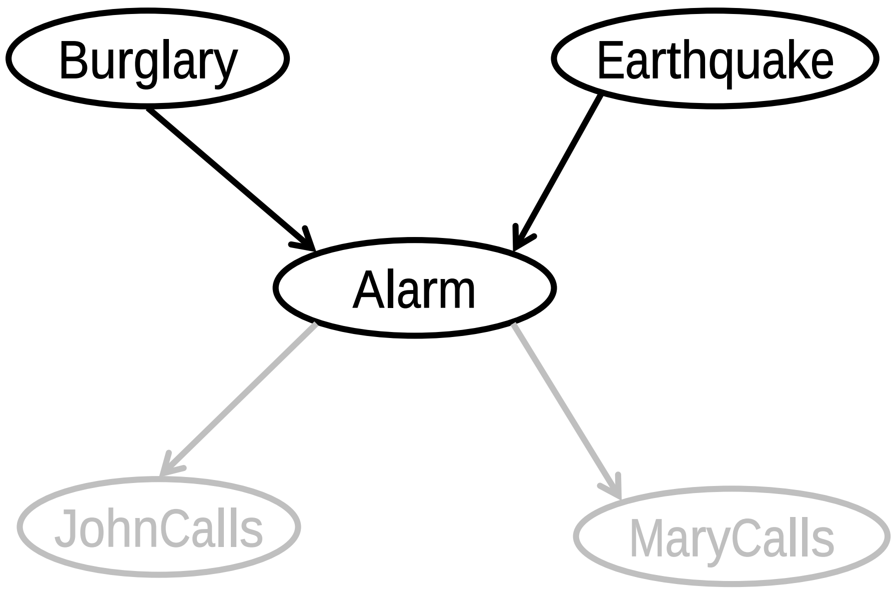
A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

The joint probability distribution

- A generic entry in the joint probability distribution $P(x_1, \ldots, x_n)$ is given by:

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



B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

The joint probability distribution

- Probability of the event that the alarm has sounded but neither a burglary nor an earthquake has occurred, and both Mary and John call:

$$P(J \land M \land A \land \neg B \land \neg E)$$
$$= P(J \mid A)\ P(M \mid A)\ P(A \mid \neg B \land \neg E)\ P(\neg B)\ P(\neg E)$$
$$= 0.9 \text{ X } 0.7 \text{ X } 0.001 \text{ X } 0.999 \text{ X } 0.998$$
$$= 0.00062$$

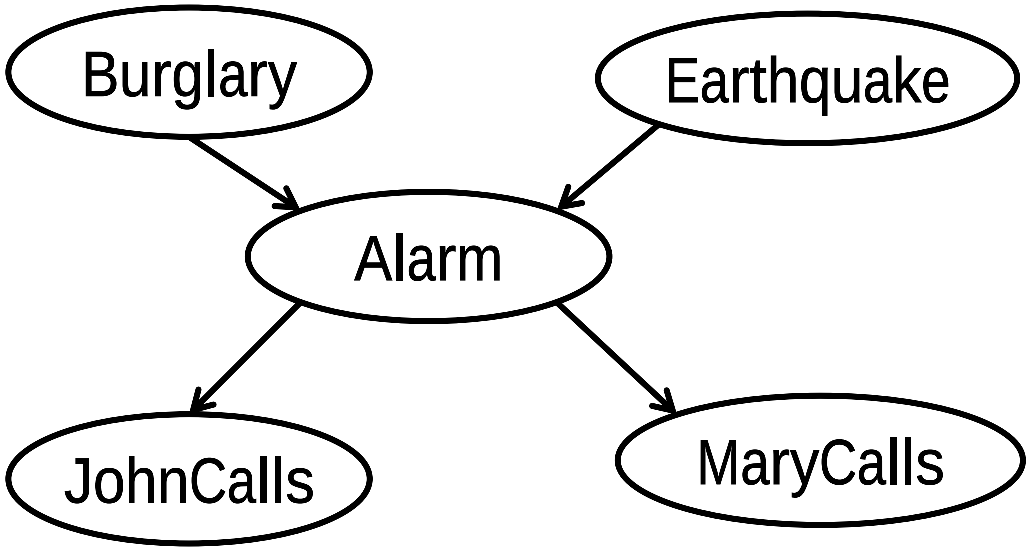
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

P(B)
0.001



The joint probability distribution

- Computation of the probabilities of several different event combinations of the Burglary-Alarm belief network example:

$P(B) = 0.001$
$P(B') = 1 – P(B) = 0.999$

$P(E) = 0.002$
$P(E') = 1 – P(E) = 0.998$

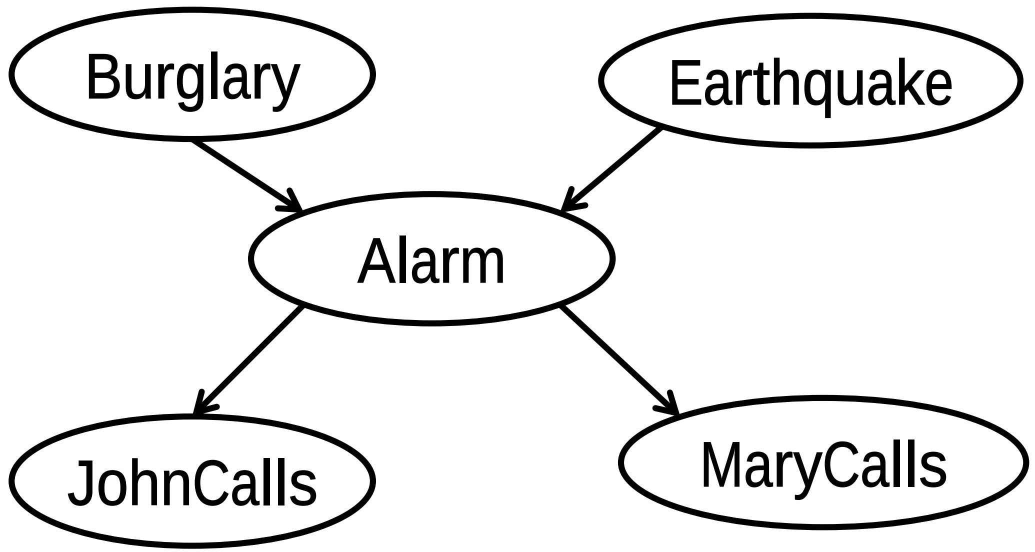
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

P(B)
0.001



The joint probability distribution

- Computation of the probabilities of several different event combinations of the Burglary-Alarm belief network example:

P(A) = P(AB’E’) + P(AB’E) + P(ABE’) + P(ABE)

= P(A | B’E’).P(B’E’) + P(A | B’E).P(B’E) + P(A | BE’).P(BE’) + P(A | BE).P(BE)

= 0.001 x 0.999 x 0.998 + 0.29 x 0.999 x 0.002 + 0.95 x 0.001 x 0.998 + 0.95 x 0.001 x 0.002

= 0.001 + 0.0006 + 0.0009 = 0.0025

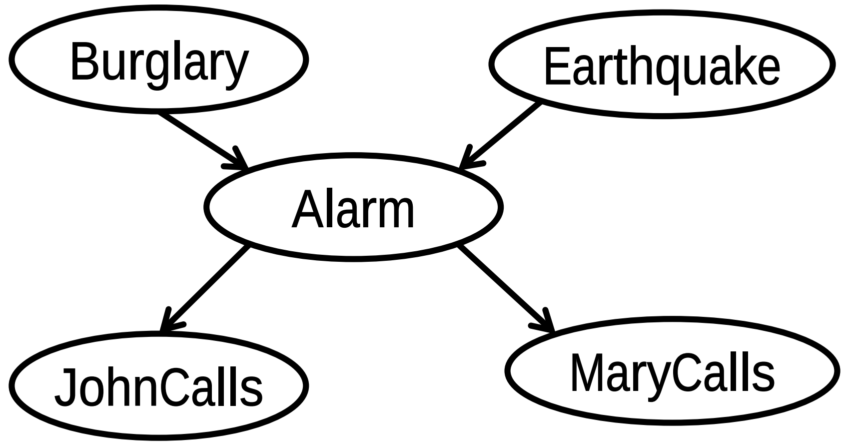
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

P(B)
0.001



The joint probability distribution: *Find* P(J)

$$P(J) = P(JA) + P(JA')$$

$$= P(J \mid A).P(A) + P(J \mid A').P(A')$$

$$= 0.9 \times 0.0025 + 0.05 \times (1 - 0.0025)$$

$$= 0.052125$$

$$P(AB) = P(ABE) + P(ABE') = 0.95 \times 0.001 \times 0.002 + 0.95 \times 0.001 \times 0.998$$

$$= 0.00095$$

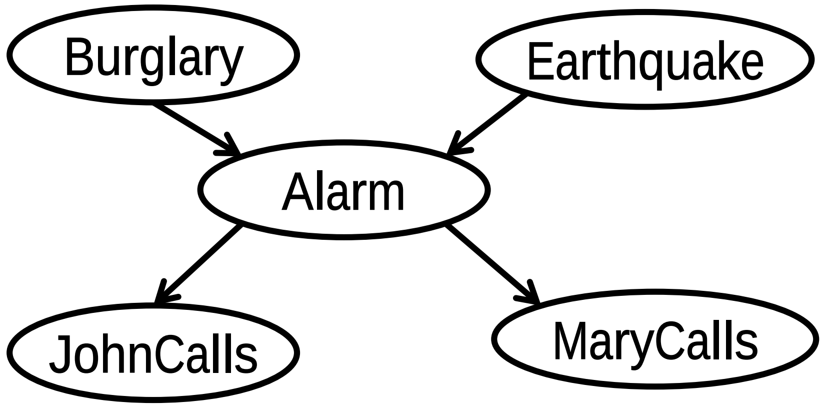
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

P(B)
0.001



The joint probability distribution: *Find* P(A'B) *and* P(AE)

$$P(A'B) = P(A'BE) + P(A'BE')$$
$$= P(A' \mid BE).P(BE) + P(A' \mid BE').P(BE')$$
$$= (1 - 0.95) \times 0.001 \times 0.002$$
$$+ (1 - 0.95) \times 0.001 \times 0.998$$
$$= 0.00005$$

$$P(AE) = P(AEB) + P(AEB')$$
$$= 0.95 \times 0.001 \times 0.002 + 0.29 \times 0.999 \times 0.002 = 0.00058$$

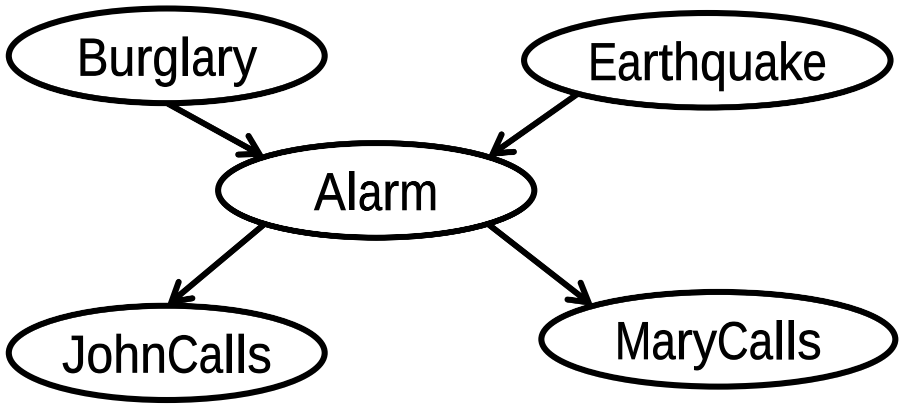
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

P(B)
0.001



The joint probability distribution

P(AE') = P(AE'B) + P(AE'B')

= 0.95 x 0.001 x 0.998 + 0.001 x 0.999 x 0.998

= 0.001945

P(A'E') = P(A'E'B) + P(A'E'B')

= P(A' | BE').P(BE') + P(A' | B'E').P(B'E')

= (1 – 0.95) x 0.001 x 0.998 + (1 – 0.001) x 0.999 x 0.998 = 0.996

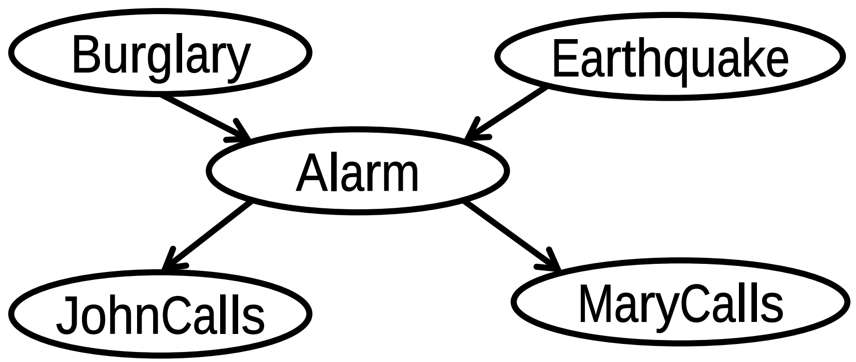
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

P(B)
0.001



The joint probability distribution: *Find* P(JB)

$$P(JB) = P(JBA) + P(JBA')$$
$$= P(J \mid AB).P(AB) + P(J \mid A'B).P(A'B)$$
$$= P(J \mid A).P(AB) + P(J \mid A').P(A'B)$$
$$= 0.9 \times 0.00095 + 0.05 \times 0.00005$$
$$= 0.00086$$

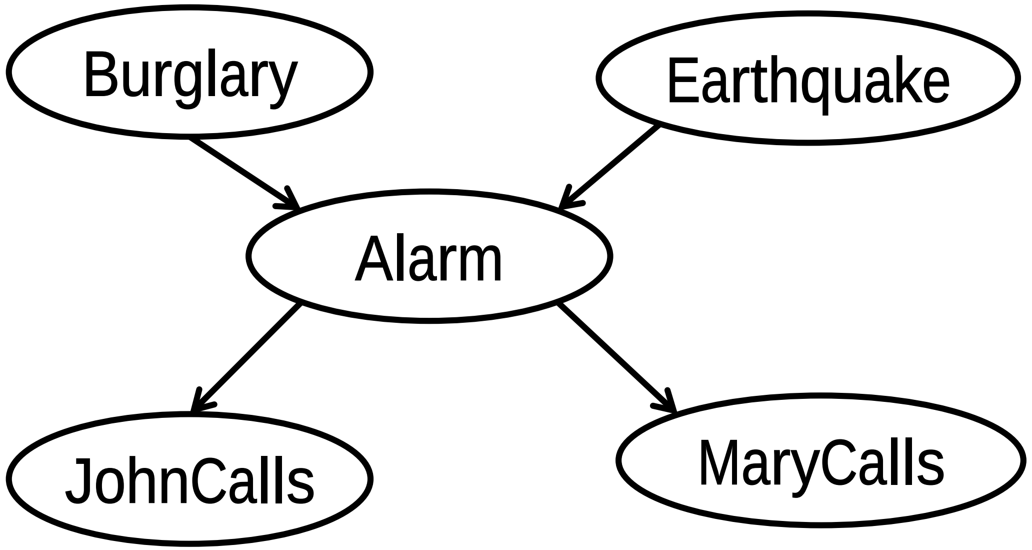
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

P(B)
0.001



The joint probability distribution

- Computation of the probabilities of several different event combinations of the Burglary-Alarm belief network example:

P(J | B) = P(JB) / P(B) = 0.00086 / 0.001 = 0.86

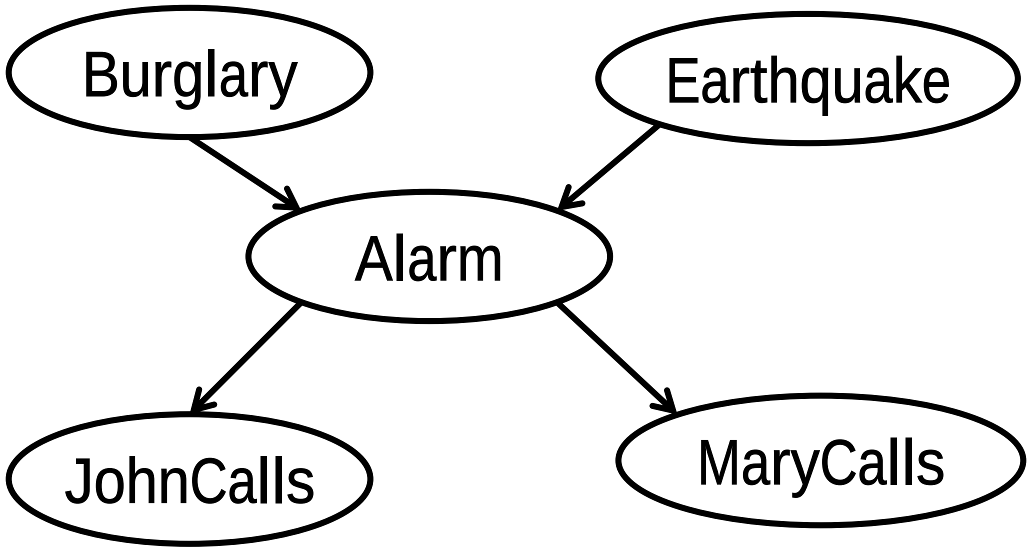
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

P(B)
0.001



The joint probability distribution

$$P(MB) = P(MBA) + P(MBA')$$
$$= P(M \mid AB).P(AB) + P(M \mid A'B).P(A'B)$$
$$= P(M \mid A).P(AB) + P(M \mid A').P(A'B)$$
$$= 0.7 \times 0.00095 + 0.01 \times 0.00005$$
$$= 0.00067$$

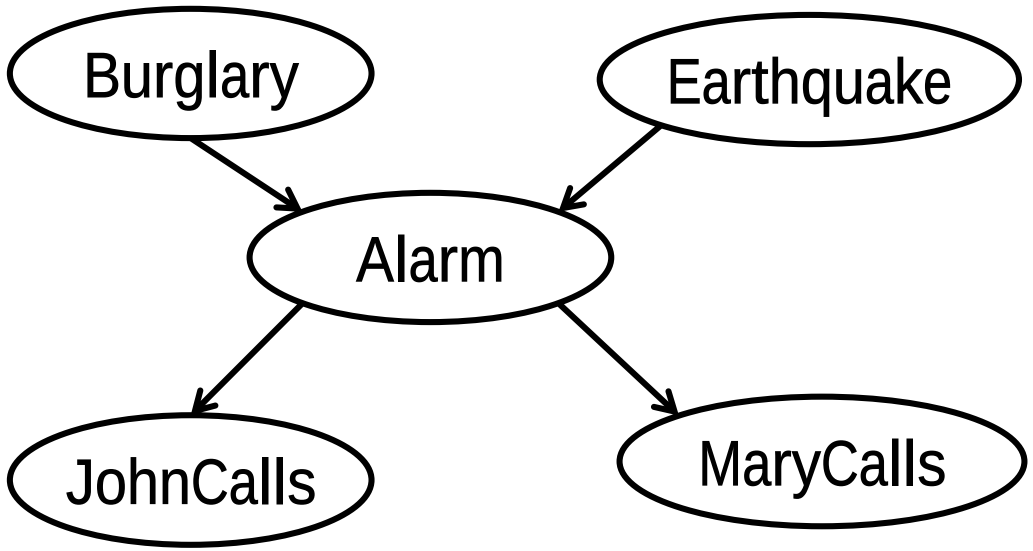
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

P(B)
0.001



The joint probability distribution

P(M | B) = P(MB) / P(B) = 0.00067 / 0.001 = 0.67

P(B | J) = P(JB) / P(J) = 0.00086 / 0.052125 = 0.016

P(B | A) = P(AB) / P(A) = 0.00095 / 0.0025 = 0.38

P(B | AE) = P(ABE) / P(AE) = [P(A | BE).P(BE)] / P(AE)

= [0.95 x 0.001 x 0.002] / 0.00058

= 0.003

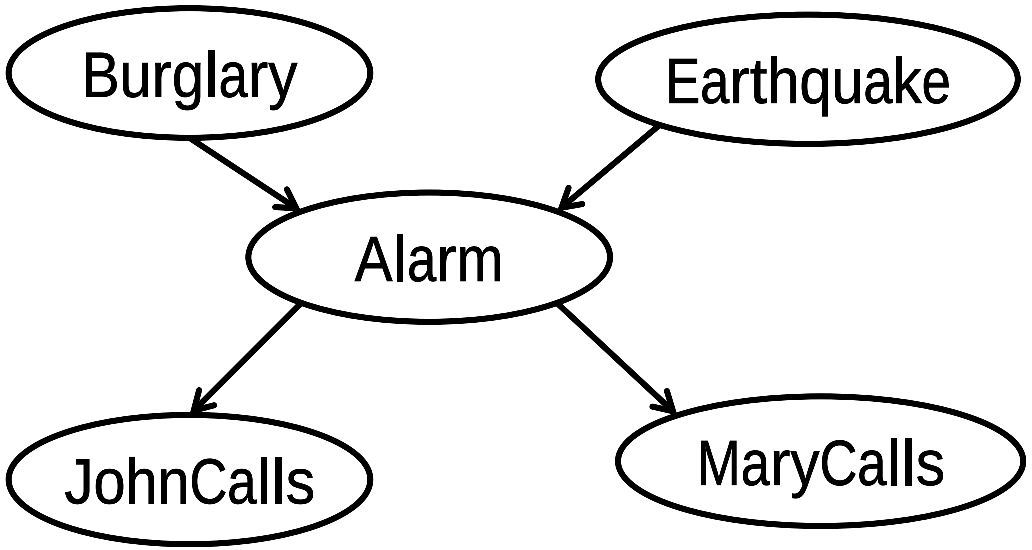
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

P(B)
0.001



The joint probability distribution

- Computation of the probabilities of several different event combinations of the Burglary-Alarm belief network example:

$$P(AJE') = P(J \mid AE').P(AE') = P(J \mid A).P(AE')$$

$$= 0.9 \times 0.001945 = 0.00175$$

$$P(A'JE') = P(J \mid A'E').P(A'E') = P(J \mid A').P(A'E')$$

$$= 0.05 \times 0.996 = 0.0498$$

$$P(JE') = P(AJE') + P(A'JE') = 0.00175 + 0.0498 = 0.05155$$

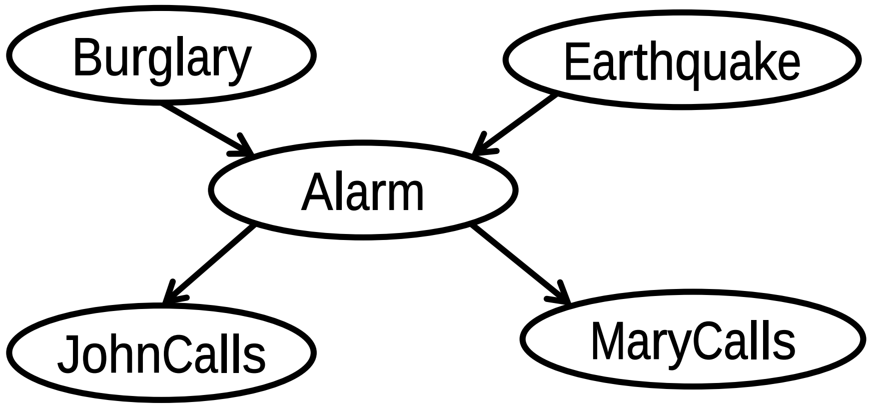
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

P(B)
0.001



The joint probability distribution

$$P(A \mid JE') = P(AJE') / P(JE') = 0.00175 / 0.05155 = 0.03$$

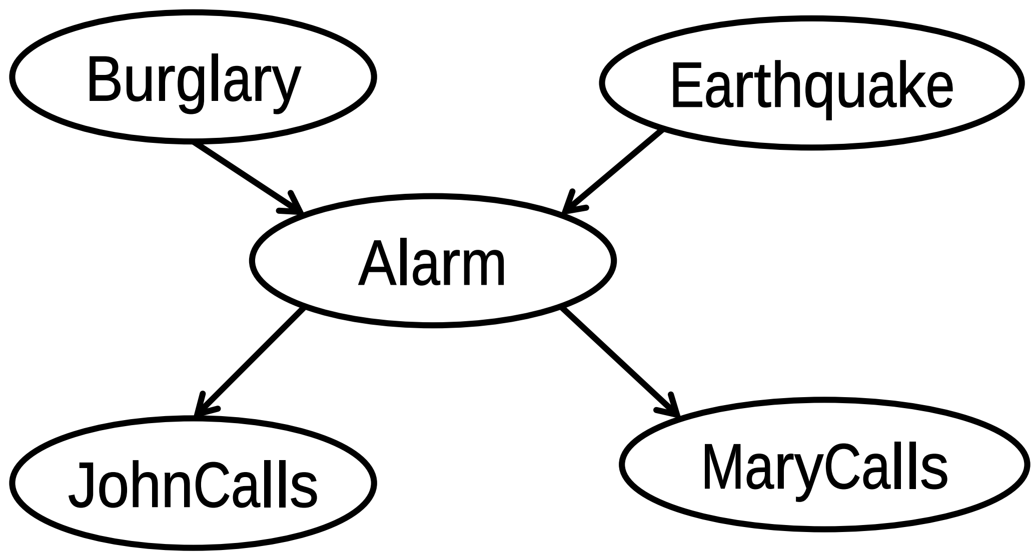
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

P(B)
0.001



The joint probability distribution

P(BJE') = P(BJE'A) + P(BJE'A')

= P(J | ABE').P(ABE') + P(J | A'BE').P(A'BE')

= P(J | A).P(ABE') + P(J | A').P(A'BE')

= 0.9 x 0.95 x 0.001 x 0.998 + 0.05 x (1 – 0.95) x 0.001 x 0.998

= 0.000856

P(B | JE') = P(BJE') / P(JE') = 0.000856 / 0.05155 = 0.017

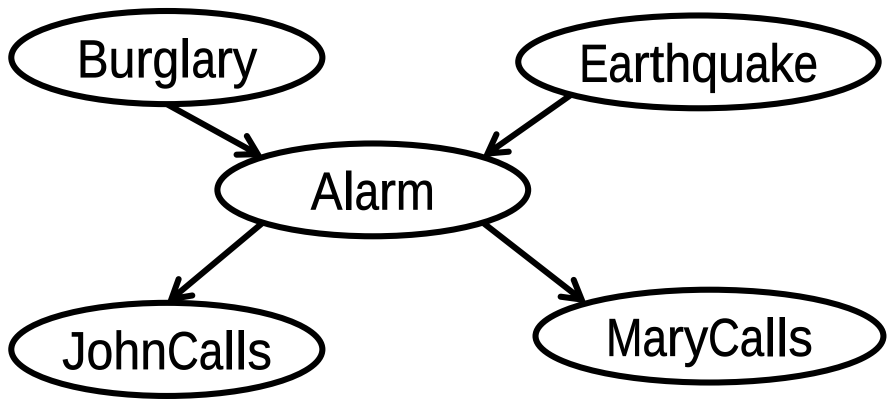
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

A	P(J)
T	0.90
F	0.05

A	P(M)
T	0.70
F	0.01

P(E)
0.002

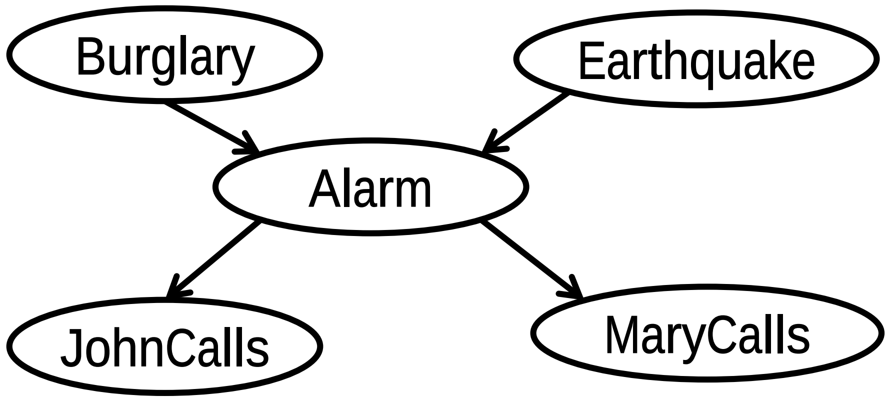
P(B)
0.001



Inferences using belief networks

- Diagnostic inferences (from effects to causes)
    - Given that JohnCalls, infer that

      P(Burglary | JohnCalls) = 0.016



- Causal inferences (from causes to effects)
    - Given Burglary, infer that

      P(JohnCalls | Burglary) = 0.86

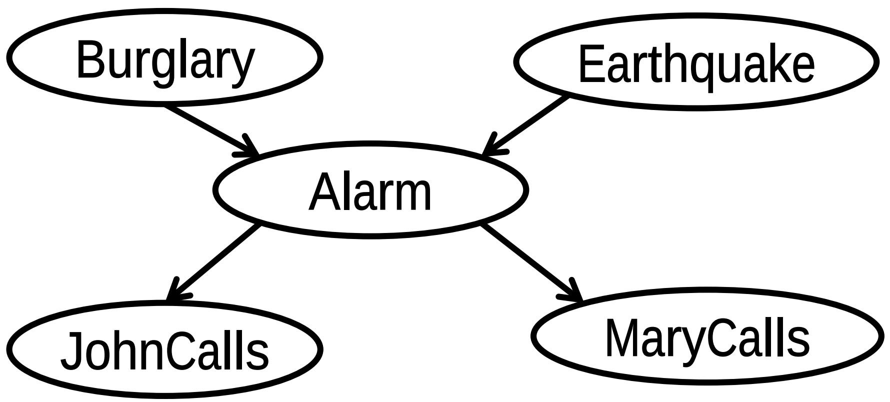
      P(MaryCalls | Burglary) = 0.67

Inferences using belief networks

- Inter-causal inferences (between causes of a 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 the cause Earthquake to false gives

    P(Alarm | JohnCalls ∧ ¬ Earthquake) = 0.003



Exercise

Three candidates run for an election as a major in a city.

According to a public opinion poll, their chances to win are 0.25, 0.35 und 0.40.

The chances that they build a bridge after they have been elected are 0.60, 0.90 and 0.80.

What is the probability that the bridge will be built after the election?

Solution: Let $C, c \in \{1, 2, 3\}$, be the random variable indicating the winning candidate and $B, b \in \{t, f\}$, the random variable indicating whether the bridge will be built. Then the total probability that the bridge will be built is

$$P(B = t) = \sum_{c=1}^{3} P(B = t|c)P(c) = 0.60 \times 0.25 + 0.90 \times 0.35 + 0.80 \times 0.40 = 0.785\,.$$

Exercise

On an airport all passengers are checked carefully.

Let T with t ∈ {0, 1} be the random variable indicating whether somebody is a terrorist (t = 1) or not (t = 0).

Let A with a ∈ {0, 1} be the variable indicating arrest.

A terrorist shall be arrested with probability P(A = 1|T = 1) = 0.98, a non-terrorist with probability P(A = 1|T = 0) = 0.001.

One in a lakh passengers is a terrorist, P(T = 1) = 0.00001.

What is the probability that an arrested person actually is a terrorist?

Solution: This can be solved directly with the Bayesian theorem.

$$
\begin{aligned}
P(T=1|A=1) &= \frac{P(A=1|T=1)P(T=1)}{P(A=1)} && (1)\\
&= \frac{P(A=1|T=1)P(T=1)}{P(A=1|T=1)P(T=1)+P(A=1|T=0)P(T=0)} && (2)\\
&= \frac{0.98 \times 0.00001}{0.98 \times 0.00001 + 0.001 \times (1-0.00001)} = 0.0097 && (3)\\
&\approx \frac{0.00001}{0.001} = 0.01 && (4)
\end{aligned}
$$

It is interesting that even though for any passenger it can be decided with high reliability (98% and 99.9%) whether (s)he is a terrorist or not, if somebody gets arrested as a terrorist, (s)he is still most likely not a terrorist (with a probability of 99%).