Consider the following KB:

1. **John likes all kinds of food.**
2. **Apples are food.**
3. **Chicken is food.**
4. **Anything anyone eats and isn't killed by is food.**
5. **Bill eats peanuts and is still alive.**
6. **Sue eats everything Bill eats.**

    i. Translate these sentences into formulas in FOPC.
    ii. Convert the formulas into clause form.
    iii. Use resolution to prove that John likes peanuts.
    iv. Use resolution to answer the question, "What food does Sue eat?"

(a) Translate these sentences into WFFs in FOPL.

1. $\forall x\ Food(x) \rightarrow Likes(John, x)$
2. $Food(Apples)$
3. $Food(Chicken)$
4. $\forall y \exists x\ Eats(y, x) \land \neg KilledBy(y, x) \rightarrow Food(x)$
5. $Eats(Bill, Peanuts) \land \neg KilledBy(Bill, Peanuts)$
6. $\forall x\ Eats(Bill, x) \rightarrow Eats(Sue, x)$

OR

a) Translate these sentences into formulas in FOPC.

1. ∀x Food(x) → Likes(John, x)
2. Food(Apples)
3. Food(Chicken)
4. ∀y ∀x Eats(y, x) ∧ ¬KilledBy(y, x) → Food(x)
5. Eats(Bill, Peanuts) ∧ ¬KilledBy(Bill, Peanuts)

eats(bill, peanuts) /\ alive(bill) => [alive(bill) {after eating peanuts} → ¬KilledBy(Bill, Peanuts)]

∀x ∀y [killed-by(y, x) → ¬ alive(y)]

6. ∀x Eats(Bill, x) → Eats(Sue, x)

(c) Convert the formulas of part (a) into clause CNF form.

1. ¬Food(x) ∨ Likes(John, x)
2. Food(Apples)
3. Food(Chicken)
4. ¬Eats(y, x) ∨ KilledBy(y, x) ∨ Food(x)
5. Eats(Bill, Peanuts) **[FACTS]**
6. ¬KilledBy(Bill, Peanuts) **[FACTS]**
7. ¬Eats(Bill, x) ∨ Eats(Sue, x)

(d) Use resolution to prove that John likes peanuts.

8. ¬Likes(John, Peanuts)	[Assumption of negated conclusion]
9. ¬Food(Peanuts)	[Resolving 1 and 8]
10. ¬Eats(y, Peanuts) ∨ KilledBy(y, Peanuts)	[Resolving 4 and 9]
11. KilledBy(Bill, Peanuts)	[Resolving 5 and 10]
12. NIL	[Resolving 6 and 11]

Home Work

(e) Use resolution to answer the question, “What food does Sue eat?”

13. ¬Eats(Sue, x) negated query
14. ¬Eats(Bill, x) 7 and 13
15. ⊥ 5 and 14, unifying Peanuts/x

2. Consider the following facts:
i. Steve only likes easy courses.
ii. Science courses are hard.
iii. All the courses in the basketweaving department are easy.
iv. BK301 is a basketweaving course.
a) Convert these facts to wffs in predicate logic.
b) Use backward chaining to answer the question "What course would Steve like?"

Converting it into FOPL (First order predicate logic)
∀x : easy(x) -> likes(steve, x)
∀x : science(x) -> hard(x) **[¬easy(x)]**
∀x : basketweaving(x) -> easy(x)
basketweaving(BK301)
likes(steve , x).

(1) ~easy(x) likes(steve,x) (2) ~science(x) ~easy(x) (3) ~humanities (x) easy(x) (4) humanities(HM101) (5) ~likes(steve,x)

(2) (6) 1&5 yields resolvent ~easy(x). • (7) 3&6 yields resolvent ~humanities (x). • (8) 4&7 yields empty clause; the substitution x/HM101 is produced by the unification algorithm which says that the only wff of the form likes(steve,x) which follows from the premises is likes(steve, HM101). Thus, resolution gives us a way to find additional assumptions.

Deduction Using Propositional Logic: Example 1

Choice of Boolean Variables a, b, c, d, ... which can take values true or false.

Boolean Formulae developed using well defined connectors ~, ∧, V, →, etc, whose meaning (semantics) is given by their truth tables.

Codification of Sentences of the argument into Boolean Formulae.

Developing the Deduction Process as obtaining truth of a Combined Formula expressing the complete argument.

Determining the Truth or Validity of the formula and thereby proving or disproving the argument and Analyzing its truth under various Interpretations.

If I am the President then I am well-known. I am the President. So I am well-known

Coding: Variables

a: I am the President

b: I am well-known

Coding the sentences:

F1: $a \rightarrow b$

F2: a

G: b

The final formula for deduction: $(F1 \wedge F2) \rightarrow G$, that is:

$((a \rightarrow b) \wedge a) \rightarrow b$

Deduction Using Propositional Logic: Example 1

Boolean variables a, b, c, d, ... which can take values true or false.

Boolean formulae developed using well defined connectors ~, ∧, ∨, →, etc, whose meaning (semantics) is given by their truth tables.

Codification of sentences of the argument into Boolean Formulae.

Developing the Deduction Process as obtaining truth of a combined formula expressing the complete argument.

Determining the Truth or Validity of the formula and thereby proving or disproving the argument and Analyzing its truth under various interpretations.

If I am the President then I am well-known. I am the President. So I am well-known

Coding: Variables

a: I am the President

b: I am well-known

Coding the sentences:

F1: $a \rightarrow b$

F2: a

G: b

The final formula for deduction: $(F1 \wedge F2) \rightarrow G$, that is: $((a \rightarrow b) \wedge a) \rightarrow b$

a	b	$a \rightarrow b$	$(a \rightarrow b) \wedge a$	$((a \rightarrow b) \wedge a) \rightarrow b$
T	T	T	T	T
T	F	F	F	T
F	T	T	F	T
F	F	T	F	T

Deduction Using Propositional Logic: Example 2

Boolean variables a, b, c, d, … which can take values true or false.

Boolean formulae developed using well defined connectors ~, Λ, V, →, etc, whose meaning (semantics) is given by their truth tables.

Codification of sentences of the argument into Boolean Formulae.

Developing the Deduction Process as obtaining truth of a combined formula expressing the complete argument.

Determining the Truth or Validity of the formula and thereby proving or disproving the argument and Analyzing its truth under various interpretations.

If I am the President then I am well-known. I am not the President. So I am not well-known

Coding: Variables

a: I am the President

b: I am well-known

Coding the sentences:

F1: $a \rightarrow b$

F2: $\sim a$

G: $\sim b$

The final formula for deduction: $(F1 \land F2) \rightarrow G$, that is: $((a \rightarrow b) \land \sim a) \rightarrow \sim b$

a	b	a → b	(a → b) Λ ~a	((a → b) Λ ~a) → ~b
T	T	T	F	T
T	F	F	F	T
F	T	T	T	F
F	F	T	T	T

Wherever Mary goes, so does the Lamb. Mary goes to School. So the Lamb goes to School.

Predicate: goes(x,y) to represent x goes to y

New Connectors: $\exists$ (there exists), $\forall$(for all)

F1: $\forall x(\text{goes}(\text{Mary}, x) \rightarrow \text{goes}(\text{Lamb}, x))$

F2: goes(Mary, School)

G: goes(Lamb, School)

To prove: $(F1 \wedge F2) \rightarrow G)$ is always true

No contractors are dependable. Some engineers are contractors. Therefore some engineers are not dependable.

<u>Predicates</u>: contractor(x), dependable(x), engineer(x)

F1: $\forall$x(contractor(x) → ~dependable(x))

[Alternative: ~$\exists$x (contractor(x) $\wedge$ dependable(x))]

F2: $\exists$x(engineer(x) $\wedge$ contractor(x))

G: $\exists$x(engineer(x) $\wedge$ ~dependable(x))

To prove: (F1 $\wedge$ F2) → G) is always true

Reasoning under Uncertainty

The intelligent way to handle the unknown

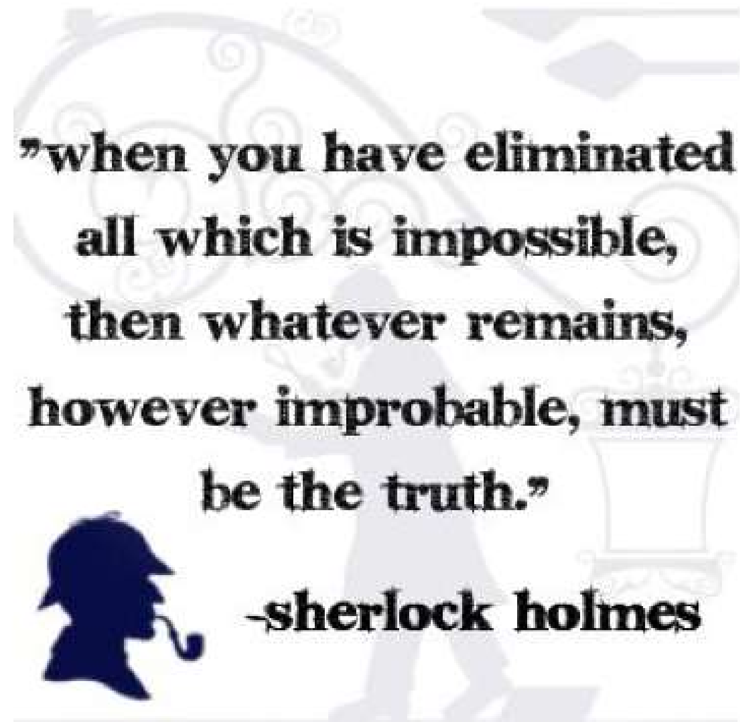
Logical Deduction versus Induction

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. $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$

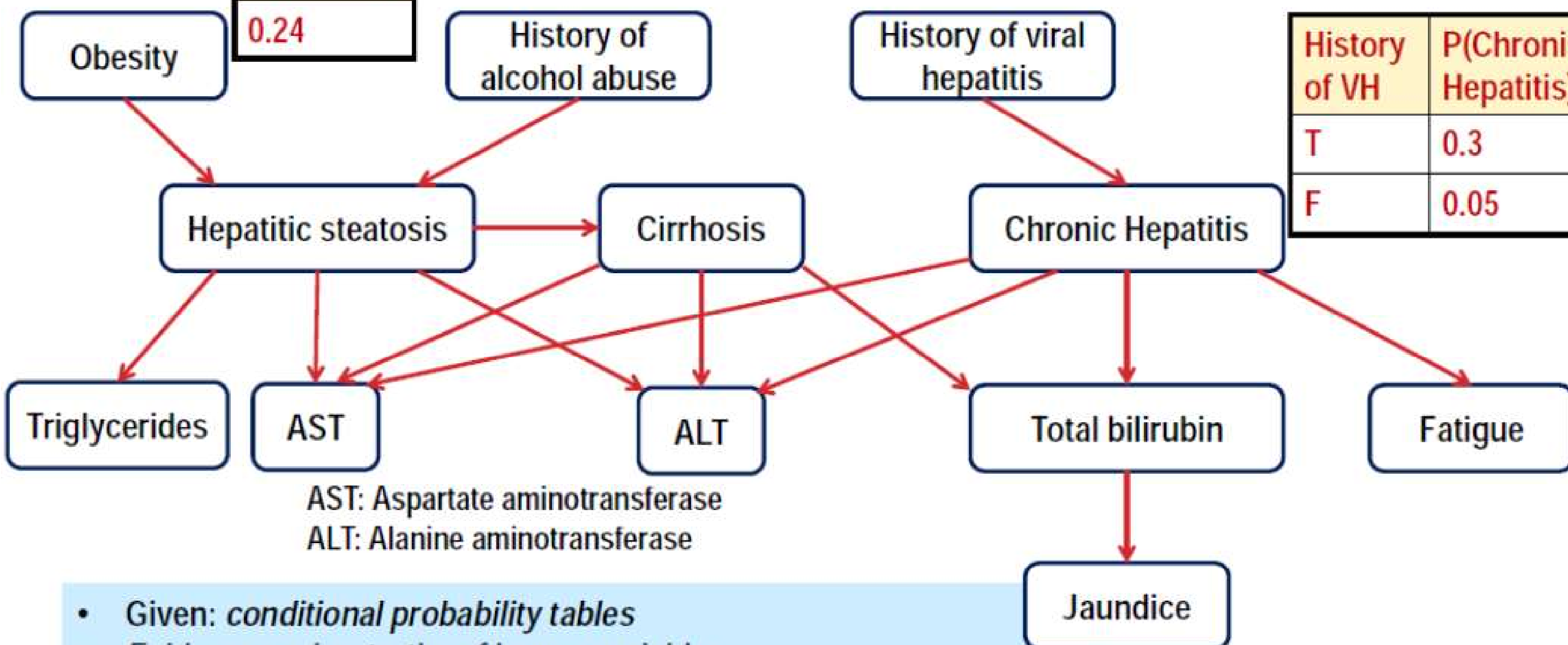
Bayes' Rule

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

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

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

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

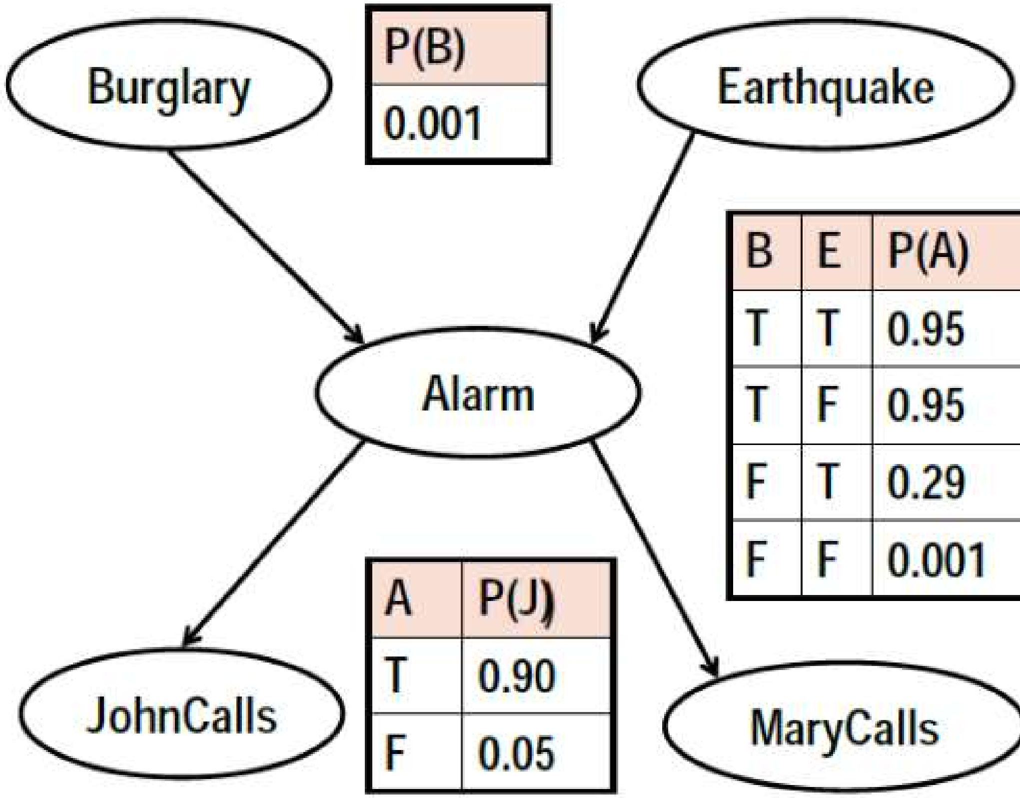
Each node has a conditional probability table that quantifies the effects that the parent have on the node.

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

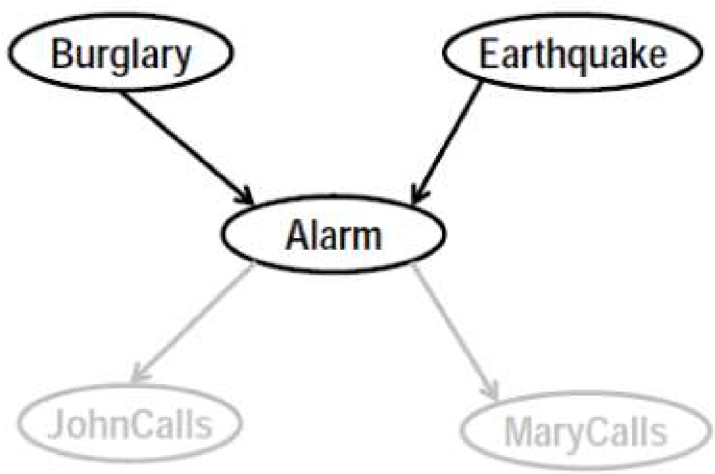
A	P(J)
T	0.90
F	0.05

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:

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

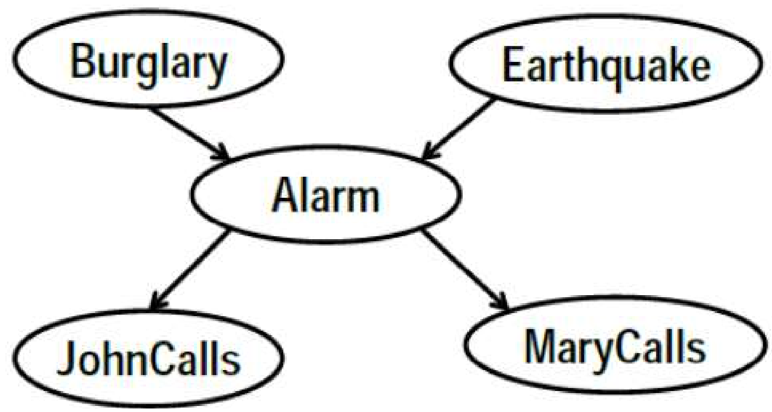


B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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:

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



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:

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

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

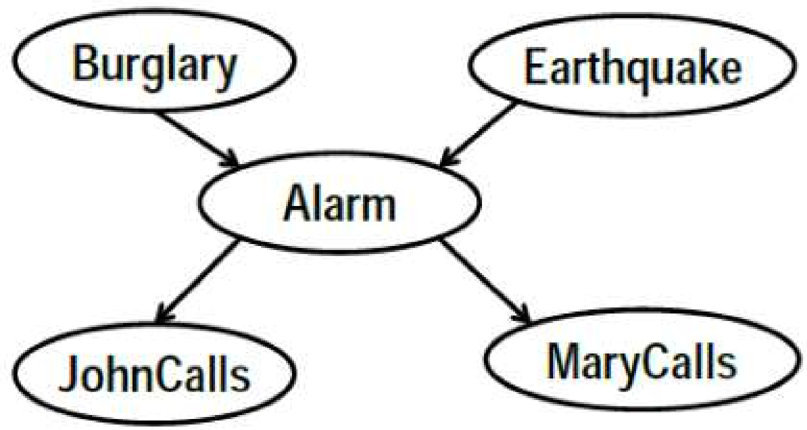
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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:

$$P(A) = P(AB'E') + P(AB'E) + P(ABE') + P(ABE)$$

$$= P(A \mid B'E').P(B'E') + P(A \mid B'E).P(B'E) + P(A \mid BE').P(BE') + P(A \mid BE).P(BE)$$

$$= 0.001 \times 0.999 \times 0.998 + 0.29 \times 0.999 \times 0.002 + 0.95 \times 0.001 \times 0.998 + 0.95 \times 0.001 \times 0.002$$

$$= 0.001 + 0.0006 + 0.0009 = 0.0025$$

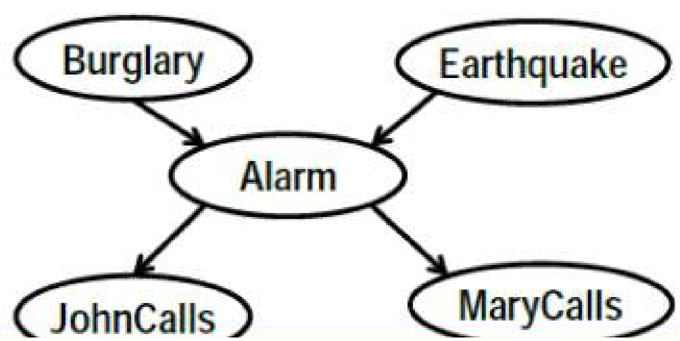
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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)

$$
\begin{aligned}
P(J) &= P(JA) + P(JA') \\
&= P(J \mid A).P(A) + P(J \mid A').P(A') \\
&= 0.9 \times 0.0025 + 0.05 \times (1 - 0.0025) \\
&= 0.052125 \\
P(AB) &= P(ABE) + P(ABE') = 0.95 \times 0.001 \times 0.002 + 0.95 \times 0.001 \times 0.998 \\
&= 0.00095
\end{aligned}
$$

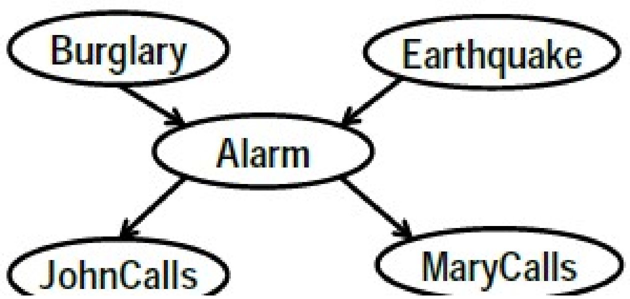
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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)

$$
\begin{aligned}
P(A'B) &= P(A'BE) + P(A'BE') \\
&= P(A' \mid BE).P(BE) + P(A' \mid BE').P(BE') \\
&= (1 - 0.95) \times 0.001 \times 0.002 \\
&+ (1 - 0.95) \times 0.001 \times 0.998 \\
&= 0.00005 \\
P(AE) &= P(AEB) + P(AEB') \\
&= 0.95 \times 0.001 \times 0.002 + 0.29 \times 0.999 \times 0.002 = 0.00058
\end{aligned}
$$

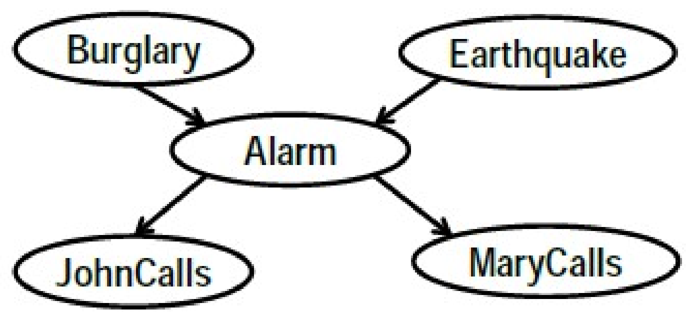
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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

$$P(AE') = P(AE'B) + P(AE'B')$$
$$= 0.95 \times 0.001 \times 0.998 + 0.001 \times 0.999 \times 0.998$$
$$= 0.001945$$

$$P(A'E') = P(A'E'B) + P(A'E'B')$$
$$= P(A' \mid BE').P(BE') + P(A' \mid B'E').P(B'E')$$
$$= (1 - 0.95) \times 0.001 \times 0.998 + (1 - 0.001) \times 0.999 \times 0.998 = 0.996$$

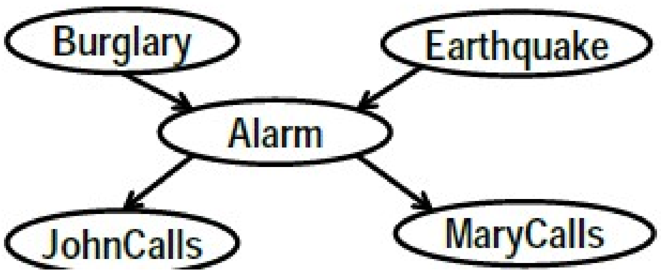
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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)

$$\begin{aligned}
P(JB) &= P(JBA) + P(JBA') \\
&= P(J \mid AB).P(AB) + P(J \mid A'B).P(A'B) \\
&= P(J \mid A).P(AB) + P(J \mid A').P(A'B) \\
&= 0.9 \times 0.00095 + 0.05 \times 0.00005 \\
&= 0.00086
\end{aligned}$$

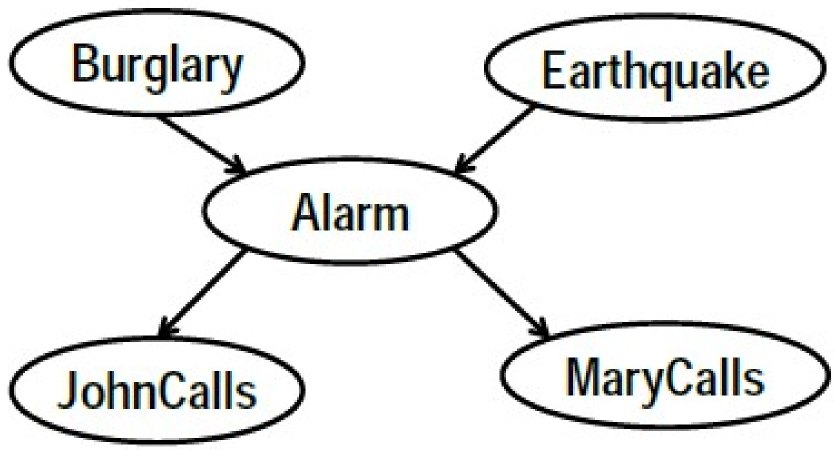
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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:

$$P(J \mid B) = P(JB) / P(B) = 0.00086 / 0.001 = 0.86$$

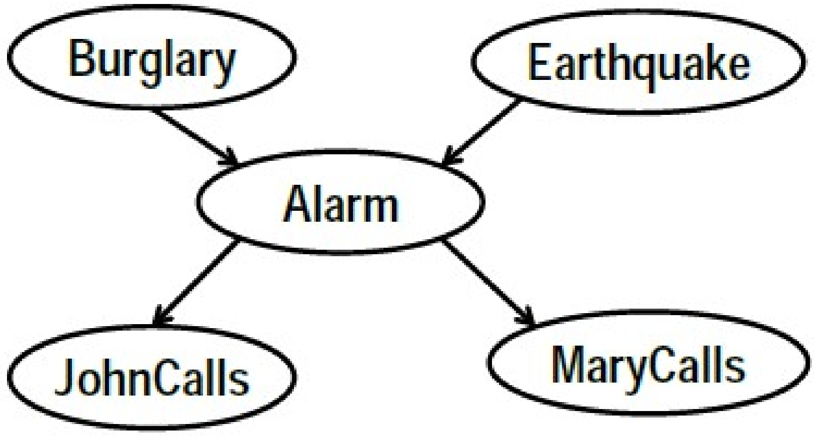
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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

$$
\begin{aligned}
P(MB) &= P(MBA) + P(MBA') \\
&= P(M \mid AB).P(AB) + P(M \mid A'B).P(A'B) \\
&= P(M \mid A).P(AB) + P(M \mid A').P(A'B) \\
&= 0.7 \times 0.00095 + 0.01 \times 0.00005 \\
&= 0.00067
\end{aligned}
$$

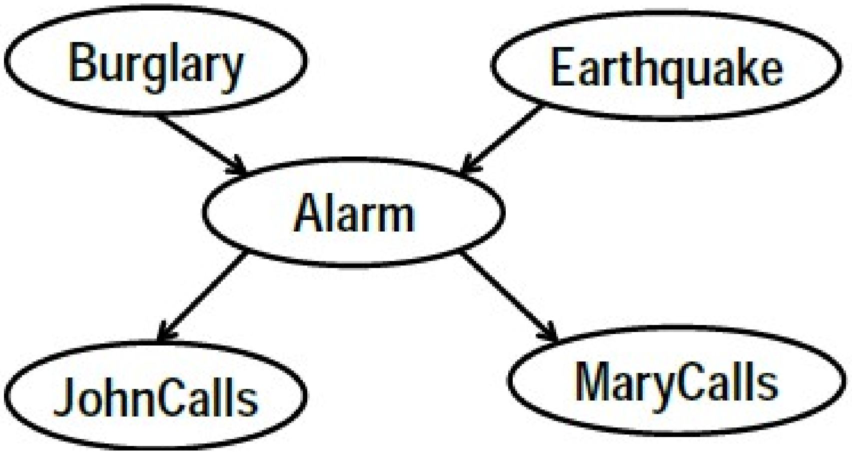
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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

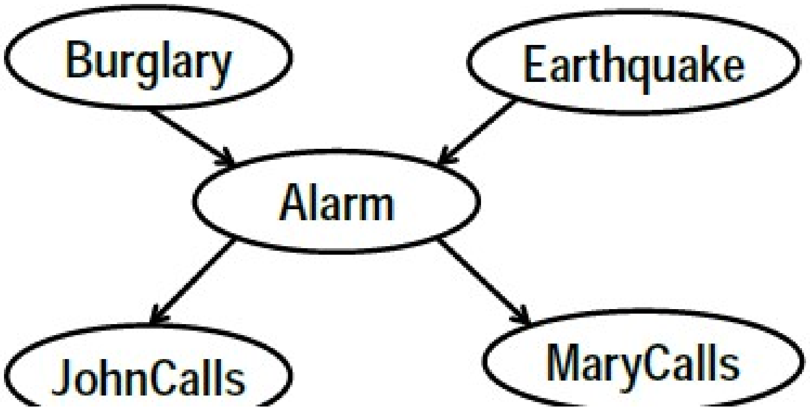
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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:

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

$$= 0.9 \times 0.001945 = 0.00175$$

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

$$= 0.05 \times 0.996 = 0.0498$$

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

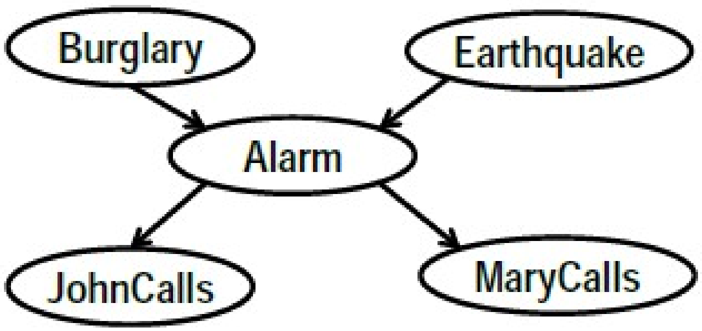
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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

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

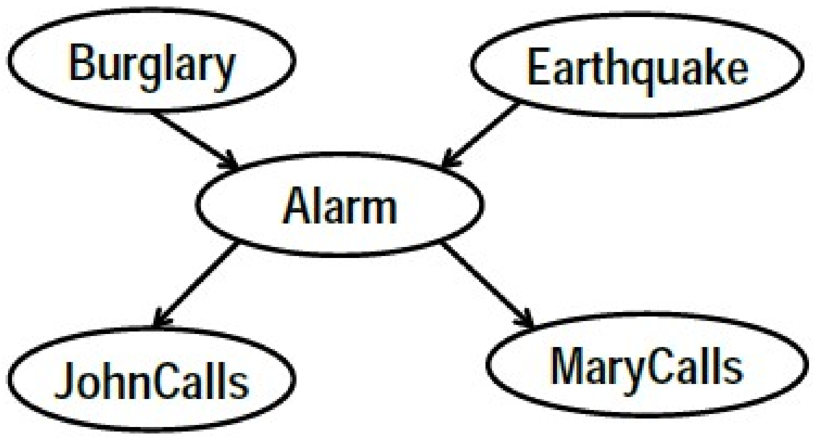
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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

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

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

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

$= 0.9 \times 0.95 \times 0.001 \times 0.998 + 0.05 \times (1 - 0.95) \times 0.001 \times 0.998$

$= 0.000856$

$P(B \mid JE') = P(BJE') / P(JE') = 0.000856 / 0.05155 = 0.017$

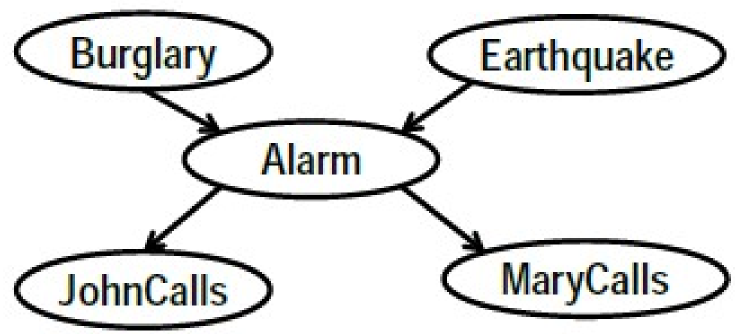
B	E	P(A)
T	T	0.95
T	F	0.95
F	T	0.29
F	F	0.001

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

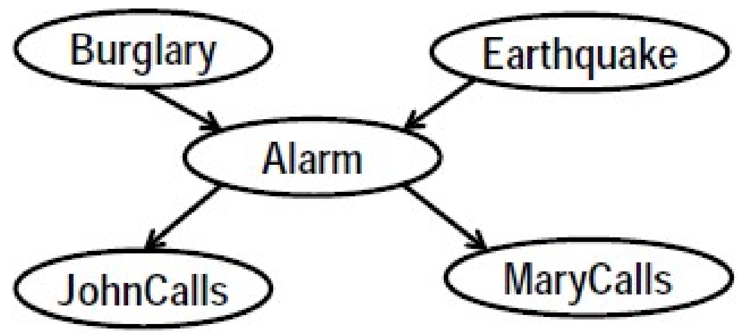


Inferences using belief networks

- 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

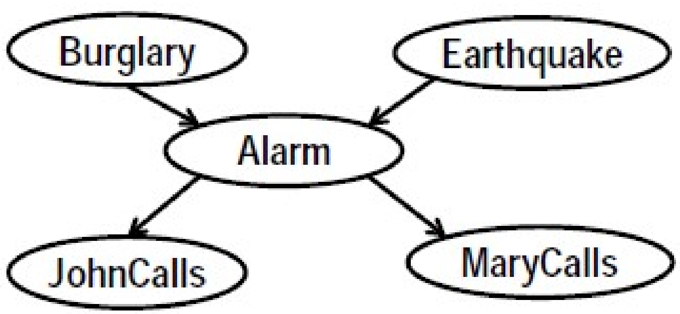
      P(MaryCalls | Burglary) = 0.67

Inferences using belief networks

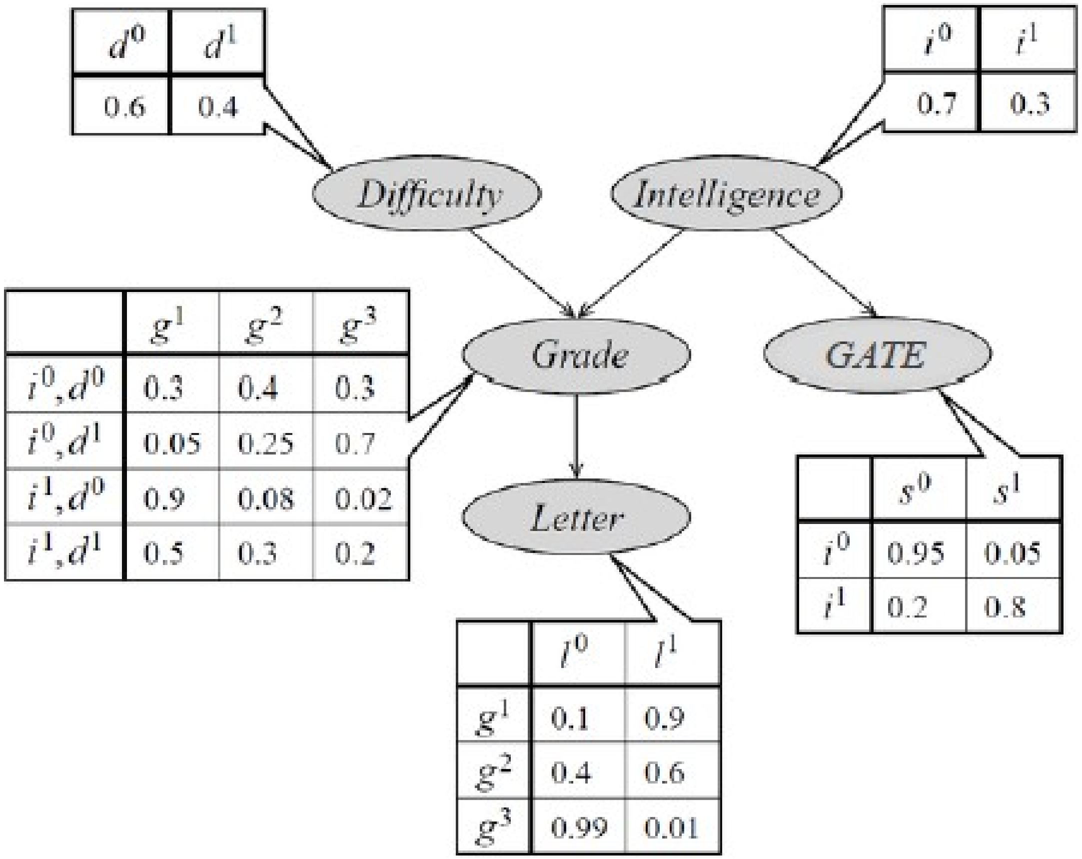
- Inter-causal inferences (between causes of a common effect)
  - Given Alarm, we have P(Burglary | Alarm) = 0.376
  - If we add evidence that Earthquake is true, then P(Burglary | Alarm $\wedge$ Earthquake) = 0.003

- Mixed inferences
  - Setting the effect JohnCalls to true and the cause Earthquake to false gives

    P(Alarm | JohnCalls $\wedge \neg$ Earthquake) = 0.003



Exercise



d^0	d^1
0.6	0.4

i^0	i^1
0.7	0.3

	g^1	g^2	g^3
i^0, d^0	0.3	0.4	0.3
i^0, d^1	0.05	0.25	0.7
i^1, d^0	0.9	0.08	0.02
i^1, d^1	0.5	0.3	0.2

	s^0	s^1
i^0	0.95	0.05
i^1	0.2	0.8

	l^0	l^1
g^1	0.1	0.9
g^2	0.4	0.6
g^3	0.99	0.01

Conditional independence

$$
\begin{aligned}
&P(x_1,\ldots,x_n) \\
&\quad = P(x_n \mid x_{n-1},\ldots,x_1)P(x_{n-1},\ldots,x_1) \\
&\quad = P(x_n \mid x_{n-1},\ldots,x_1)P(x_{n-1} \mid x_{n-2},\ldots,x_1) \\
&\qquad\qquad \ldots P(x_2 \mid x_1)P(x_1) \\
&\quad = \prod_{i=1}^{n} P(x_i \mid x_{i-1},\ldots,x_1)
\end{aligned}
$$

- The belief network represents conditional independence:

$$P(X_i \mid X_i,\ldots,X_1) = P(X_i \mid Parents(X_i))$$