Reasoning under Uncertainty

Instructors: Dr. Durgesh Singh

CSE Discipline, PDPM IIITDM, Jabalpur -482005

Reasoning under uncertainty

- Agents in the real world need to handle uncertainty, whether due to partial observability, nondeterminism, or adversaries.
- An agent may never know for sure what state it is in now or where it will end up after a sequence of actions.

Nature of Uncertain Knowledge

- Let us try to write rules for dental diagnosis using propositional logic, so that we can see how the logical approach breaks down. Consider the following simple rule:

$$\text{Toothache} \Rightarrow \text{Cavity}.$$

- The problem is that this rule is wrong.
- Not all patients with toothaches have cavities; some of them have gum disease, swelling, or one of several other problems:

$$\text{Toothache} \Rightarrow \text{Cavity} \lor \text{GumProblem} \lor \text{Swelling} \lor \ldots\ldots$$

Nature of Uncertain Knowledge

- In order to make the rule true, we have to add an almost unlimited list of possible problems. We could try turning the rule into a causal rule:

$$\text{Cavity} \Rightarrow \text{Toothache}$$

But this rule is also not right; not all cavities cause pain. Toothache and a Cavity are always not connected, so the judgement may go wrong.

Nature of Uncertain Knowledge

- This is typical of the medical domain, as well as most other judgmental domains: law, business, design, automobile repair, gardening, dating, and so on.
- The agent's knowledge can at best provide only a degree of belief in the relevant sentences.
- Our main tool for dealing with degrees of belief is probability theory.
- A logical agent believes each sentence to be true or false or has no opinion, whereas a probabilistic agent may have a numerical degree of belief between 0 (for sentences that are certainly false) and 1 (certainly true).

Basic Probability Notation

- Random variables are typically divided into three kinds, depending on the type of the domain:
- **Boolean random variables**, such as Cavity, have the domain (true, false) or (1,0)
- **Discrete random variables**, take on values from a countable domain. For example, the domain of Weather might be (sunny, rainy, cloudy, snow).
- **Continuous random variables** (bounded or unbounded) take on values from the real numbers. Ex: $temp=21.4$; $temp<21.4$ or $temp< 1$.

Atomic events or sample points

- Atomic event: A complete specification of the state of the world about which the agent is uncertain
- E.g., if the world consists of only two Boolean variables Cavity and Toothache, then there are 4 distinct atomic events:

$$Cavity = false \wedge Toothache = false$$

$$Cavity = false \wedge Toothache = true$$

$$Cavity = true \wedge Toothache = false$$

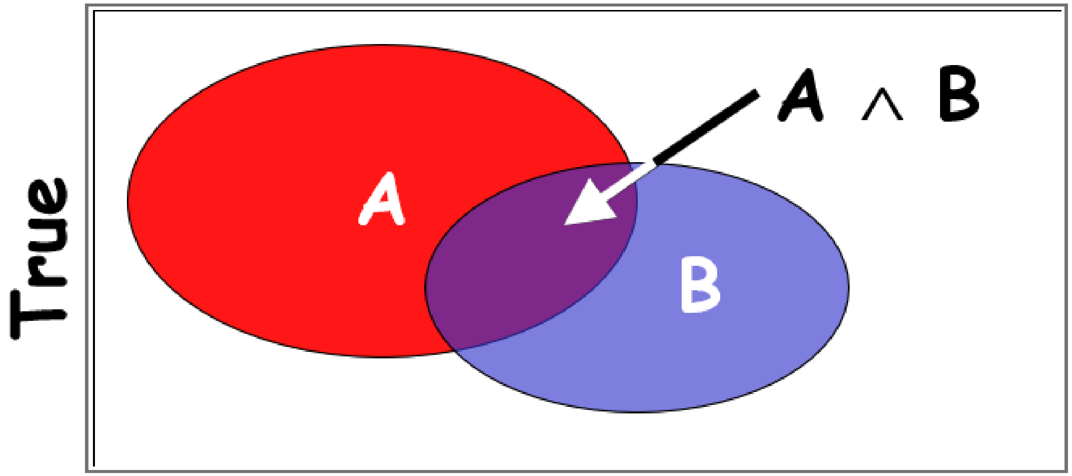
$$Cavity = true \wedge Toothache = true$$

- Atomic events are mutually exclusive and exhaustive
- When two events are mutually exclusive, it means they cannot both occur at the same time.
- When two events are exhaustive, it means that one of them must occur.

Axioms of Probability Theory

- All probabilities between 0 and 1
  - $0 \leq P(A) \leq 1$
  - $P(true) = 1$
  - $P(false) = 0$.
- The probability of disjunction is:

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



Prior probability

- The unconditional or prior probability associated with a proposition **A** is the degree of belief according to the absence of any other information;
- It is written as P (A).
- For example, if the prior probability that I have a cavity is 0.1, then we would write

P (Cavity= true) = 0.1 or P (cavity) = 0.1

- P (A) can be used only when there is no other information.
- As soon as some new information is known, we must reason with the conditional probability of a given that new information.

Prior probability…

- Sometimes, we will want to talk about the probabilities of all the possible values of a random variable.
- In that case, we will use an expression such as P(Weather), which denotes a vector of values for the probabilities of each individual state of the weather.
- Instead of writing these four equations

P (Weather = sunny) = 0.7

P (Weather= rain) = 0.2

P (Weather= cloudy) = 0.08

P(Weather = snow) = 0.02

we may simply write: P(Weather) = (0.7,0.2,0.08,0.02) (Note that the probabilities sum to 1)

- This statement defines a prior probability distribution for the random variable Weather.

Prior probability…

- Joint probability distribution for a set of random variables gives the probability of every atomic event on those random variables
- P(Weather, Cavity) = a 4 × 2 matrix of values:

Weather =	sunny	rainy	cloudy	snow
Cavity = true	0.144	0.02	0.016	0.02
Cavity = false	0.576	0.08	0.064	0.08

- A full joint distribution specifies the probability of every atomic event and is therefore a complete specification of one's uncertainty about the world in question.

Conditional or posterior probability

- The notation used is P(a l b),where a and b are any proposition.

This is read as "the probability of a, given that all we know is b." For example,

P(cavity l toothache) = 0.8

“indicates that if a patient is observed to have a toothache and no other information is yet available, then the probability of the patient's having a cavity will be 0.8.”

Conditional or posterior probability

- Conditional probabilities can be defined in terms of unconditional probabilities.

$$P(a|b) = \frac{P(a \wedge b)}{P(b)}$$

holds whenever $P(b)>0$

This equation can be written as

$P(a \wedge b) = P(a|b) * P(b)$ (which is called product rule)

Alternative way:

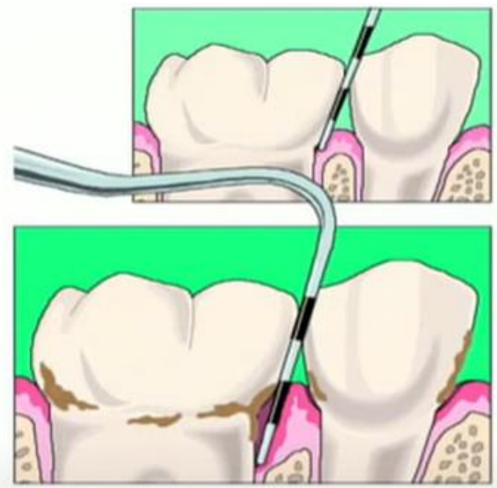
$P(a \wedge b) = P(b|a) * P(a)$

Chain Rule/Product Rule

$$P(X_1, \ldots, X_n) = P(X_n \mid X_1..X_{n-1}) P(X_{n-1} \mid X_1..X_{n-2}) \ldots P(X_1)$$
$$= \prod P(X_i \mid X_1, ..X_{i-1})$$

Example

A domain consisting of just the three Boolean variables $Toothache$, $Cavity$, and $Catch$ (the dentist's nasty steel probe catches in my tooth).



Inference Using Full Joint Distributions

Start with the joint distribution:

	toothache		*¬ toothache*	
	catch	*¬ catch*	*catch*	*¬ catch*
cavity	**.108**	**.012**	**.072**	**.008**
¬ cavity	**.016**	**.064**	**.144**	**.576**

For any proposition ϕ, sum the atomic events where it is true:

$$P(\phi) = \Sigma_{\omega:\omega \models \phi} P(\omega)$$

P(toothache)=.108+.012+.016+.064

= .20 or 20%

Inference Using Full Joint Distributions

Start with the joint distribution:

	toothache		$\neg$ *toothache*	
	catch	$\neg$ *catch*	*catch*	$\neg$ *catch*
cavity	.108	.012	.072	.008
$\neg$ *cavity*	.016	.064	.144	.576

For any proposition ϕ, sum the atomic events where it is true:

$$P(\phi) = \Sigma_{\omega:\omega \models \phi} P(\omega)$$

$P(toothache \vee cavity)$ = .20 + .072 + .008

.28

Inference Using Full Joint Distributions

Start with the joint distribution:

	toothache		¬ *toothache*	
	catch	¬ *catch*	*catch*	¬ *catch*
cavity	.108	.012	.072	.008
¬ *cavity*	.016	.064	.144	.576

Can also compute conditional probabilities:

$$P(\neg cavity | toothache) = \frac{P(\neg cavity \wedge toothache)}{P(toothache)}$$

$$= \frac{0.016 + 0.064}{0.108 + 0.012 + 0.016 + 0.064} = 0.4$$

Problems with joint distribution ??

- Worst case time: $O(d^n)$
  - Where d = max arity
  - And n = number of random variables
- Space complexity also $O(d^n)$
  - Size of joint distribution

Independence

- *A* and *B* are *independent* iff:

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

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

Therefore, if A and B are independent:

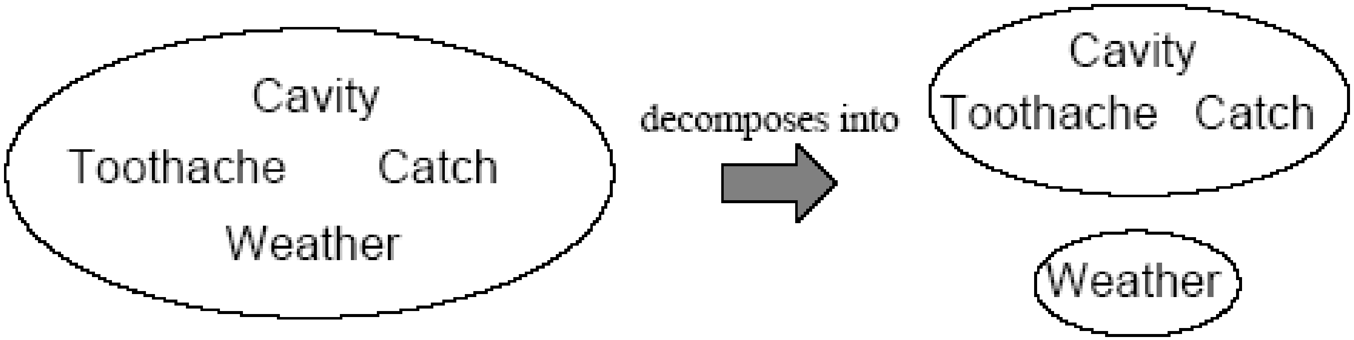
$$P(A \mid B) = \frac{P(A \wedge B)}{P(B)} = P(A)$$

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

Independence…

A and B are independent iff

$$\mathbf{P}(A|B) = \mathbf{P}(A) \quad \text{or} \quad \mathbf{P}(B|A) = \mathbf{P}(B) \quad \text{or} \quad \mathbf{P}(A,B) = \mathbf{P}(A)\mathbf{P}(B)$$



$$\mathbf{P}(Toothache, Catch, Cavity, Weather)$$
$$= \mathbf{P}(Toothache, Catch, Cavity)\mathbf{P}(Weather)$$

32 entries reduced to 12;

Complete independence is powerful but rare. What to do if it doesn't hold?

Conditional Independence

$\mathbf{P}(Toothache, Cavity, Catch)$ has $2^3 - 1 = 7$ independent entries

If I have a cavity, the probability that the probe catches in it doesn't depend on whether I have a toothache:

(1) $P(catch|toothache, cavity) = P(catch|cavity)$

The same independence holds if I haven't got a cavity:

(2) $P(catch|toothache, \neg cavity) = P(catch|\neg cavity)$

$Catch$ is *conditionally independent* of $Toothache$ given $Cavity$:

$\mathbf{P}(Catch|Toothache, Cavity) = \mathbf{P}(Catch|Cavity)$

Conditional Independence

- The general definition of **conditional independence** of two variables X and Y, given a third variable Z is

(I) $$\mathbf{P}(X, Y \mid Z) = \mathbf{P}(X \mid Z)\mathbf{P}(Y \mid Z).$$

(II) $$\mathbf{P}(X \mid Y, Z) = \mathbf{P}(X \mid Z) \quad \text{and} \quad \mathbf{P}(Y \mid X, Z) = \mathbf{P}(Y \mid Z)$$

Conditional Independence II

P(catch | toothache, cavity) = P(catch | cavity)

P(catch | toothache, ¬cavity) = P(catch | ¬cavity)

Equivalent statements:

$$\mathbf{P}(Toothache|Catch, Cavity) = \mathbf{P}(Toothache|Cavity)$$

$$\mathbf{P}(Toothache, Catch|Cavity) = \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)$$

Write out full joint distribution using chain rule:

$$\begin{aligned}
&\mathbf{P}(Toothache, Catch, Cavity) \\
&= \mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch, Cavity) \\
&= \mathbf{P}(Toothache|Catch, Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity) \\
&= \mathbf{P}(Toothache|Cavity)\mathbf{P}(Catch|Cavity)\mathbf{P}(Cavity)
\end{aligned}$$

I.e., $2 + 2 + 1 = 5$ independent numbers (equations 1 and 2 remove 2)

Power of Cond. Independence

- Often, using conditional independence reduces the storage complexity of the joint distribution from exponential to linear!!

- Conditional independence is the most basic & robust form of knowledge about uncertain environments.

Bayes Rule

$$P(H \mid E) = \frac{P(E \mid H)P(H)}{P(E)}$$

Simple proof from def of conditional probability:

$$P(H \mid E) = \frac{P(H \wedge E)}{P(E)}$$ (Def. cond. prob.)

$$P(E \mid H) = \frac{P(H \wedge E)}{P(H)}$$ (Def. cond. prob.)

$$P(H \wedge E) = P(E \mid H)P(H)$$ (Mult by P(H) in line 2)

$$P(H \mid E) = \frac{P(E \mid H)P(H)}{P(E)}$$ (Substitute #3 in #1)

Use to Compute *Diagnostic* Probability from *Causal* Probability

$$P(Cause|Effect) = \frac{P(Effect|Cause)P(Cause)}{P(Effect)}$$

E.g. let M be meningitis, S be stiff neck

P(M) = 0.0001,

P(S) = 0.1,

P(S|M)= 0.8

P(M|S) = $= \frac{P(s|m)P(m)}{P(s)} = \frac{0.8 \times 0.0001}{0.1} = 0.0008$

Note: posterior probability of meningitis still very small!

Bayes Rule

- Does patient have cancer or not?

Given: A patient takes a lab test, and the result comes back positive. The test returns a correct positive result in only 98% of the cases in which the disease is present, and a correct negative result in only 97% of the cases in which the disease is not present. Furthermore, 0.008 of the entire population have this cancer.

$$P(cancer) = \qquad P(\neg cancer) =$$

$$P(+ \mid cancer) = \qquad P(- \mid cancer) =$$

$$P(+ \mid \neg cancer) = \qquad P(- \mid \neg cancer) =$$

$$P(cancer) = 0.008 \qquad P(\neg cancer) = 0.992$$

$$P(+ \mid cancer) = 0.98 \qquad P(- \mid cancer) = 0.02$$

$$P(+ \mid \neg cancer) = 0.03 \qquad P(- \mid \neg cancer) = 0.97$$

$$P(cancer \mid +) = \frac{P(+ \mid cancer) P(cancer)}{P(+)};$$

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

$$P(cancer \mid +) P(+) = 0.98 \times 0.008 = 0.0078;$$

$$P(\neg cancer \mid +) P(+) = 0.03 \times 0.992 = 0.0298$$

$$P(+) = 0.0078 + 0.0298$$

$$P(cancer \mid +) = 0.21; \qquad P(\neg cancer \mid +) = 0.79$$

The patient, more likely than not, does not have cancer

Bayesian Networks

- In general, joint distribution over set of variables ($X_{1,}$ $X_{1,}$ … , X_n) requires exponential space for representation & inference.
- We also saw that independence and conditional independence relationships among variables can greatly reduce the number of probabilities that need to be specified in order to define the full joint distribution.
- BNs(a graphical representation) is a data structure
  - represents the dependencies among variables and
  - give a concise specification of any full joint probability distribution

Chain rule in Bayesian Networks

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

The general assertion that, for every variable Xi in the Bayesian network,

$$\mathbf{P}(X_i | X_{i-1}, \ldots, X_1) = \mathbf{P}(X_i | Parents(X_i))$$

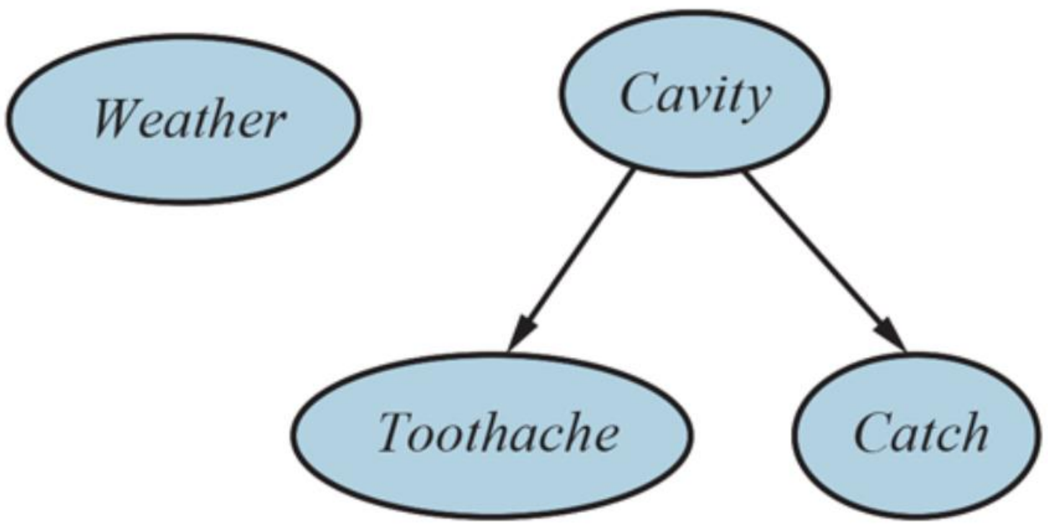
$$P(x_1, \ldots, x_n) = \prod_{i=1}^{n} P(x_i | parents(X_i)).$$

Bayes Networks

- A Bayesian network is a directed graph in which each node is annotated with quantitative probability information.
- The full specification is as follows:
  1. Each node corresponds to a random variable, which may be discrete or continuous.
  2. Directed links or arrows connect pairs of nodes. If there is an arrow from node X to node Y , X is said to be a parent of Y.
  3. Each node X_i, has a conditional probability distribution P (Xi | Parents (Xi)) that quantifies the effect of the parents on the node.
  4. The graph has no directed cycles (and hence is a directed, acyclic graph, or DAG).

Example

Topology of network encodes conditional independence assertions:



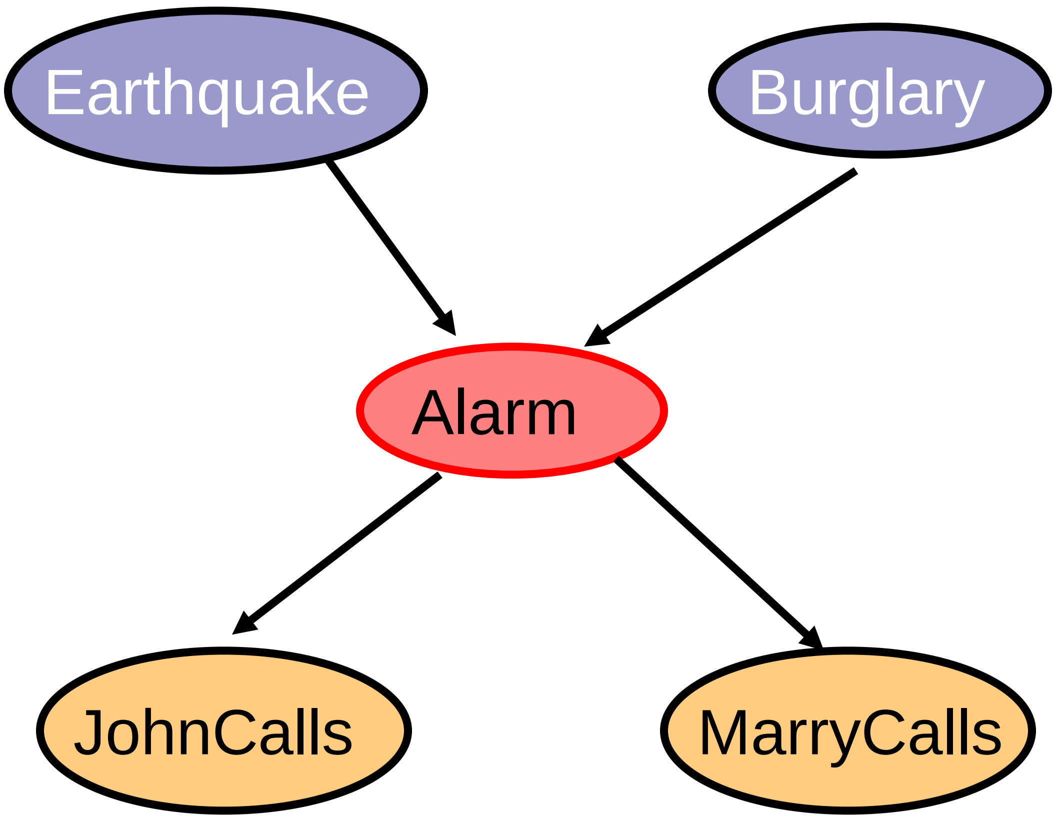
A simple Bayesian network in which $Weather$ is independent of the other three variables and $Toothache$ and $Catch$ are conditionally independent, given $Cavity$.

Example: Burglar Alarm

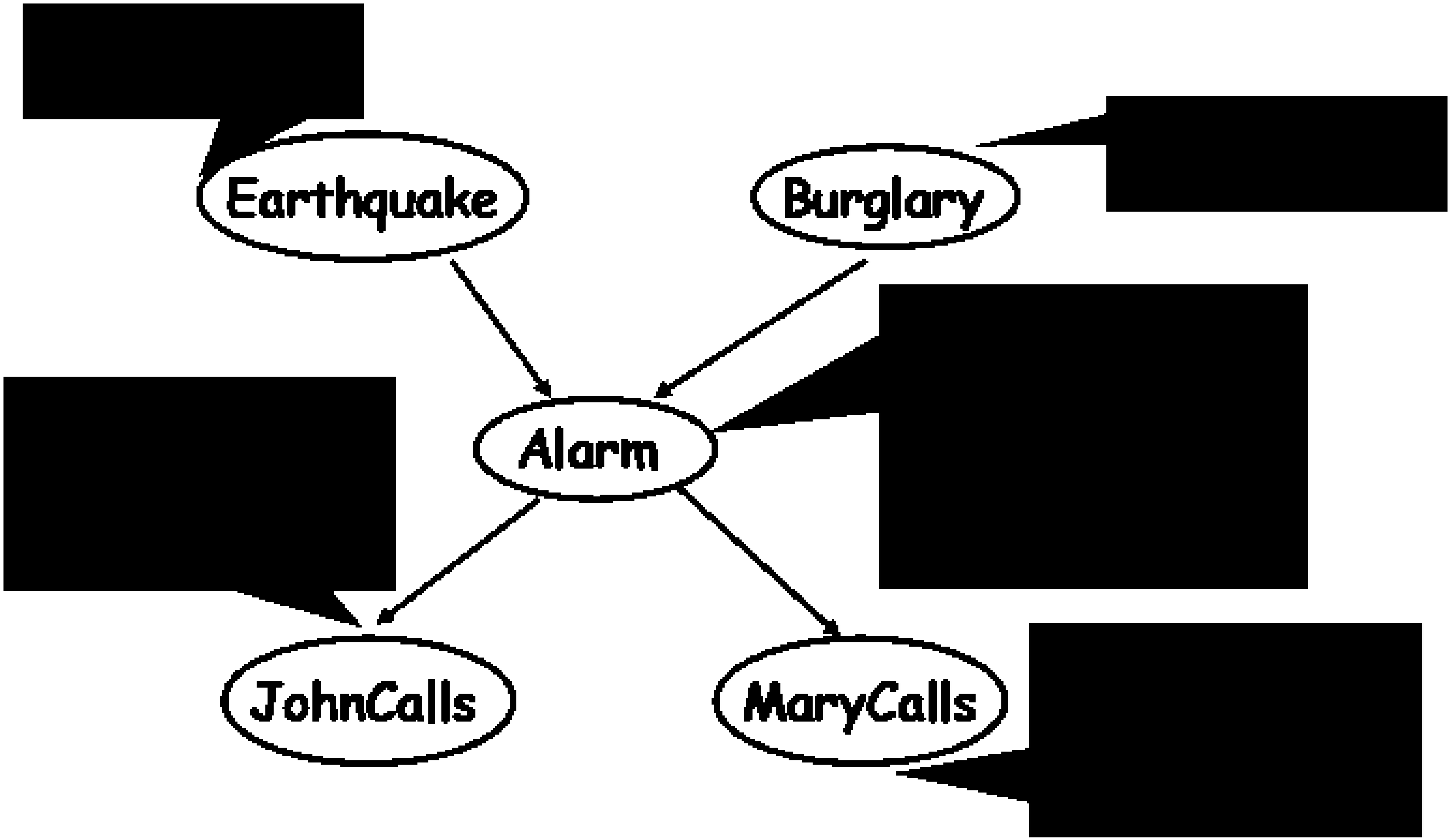
- You have a new burglar alarm installed at home.
- It is reliable at detecting a burglary, but also responds on occasion to minor earthquakes.
- You also have two neighbors, John and Mary, who have promised to call you at work when they hear the alarm.
- John always calls when he hears the alarm, but sometimes confuses the telephone ringing with the alarm and calls then, too.
- Mary, on the other hand, likes loud music and sometimes misses the alarm altogether.

Given the evidence of who has or has not called, we would like to estimate the probability of a burglary.

Example: Burglar Alarm



Example Bayes Net: Burglar Alarm



- Notice that the network does not have nodes corresponding to
    - Mary’s currently listening to loud music or
    - The telephone ringing and confusing John.
- These factors are summarized in the uncertainty associated with the links from Alarm to to JohnCalls and MaryCalls .

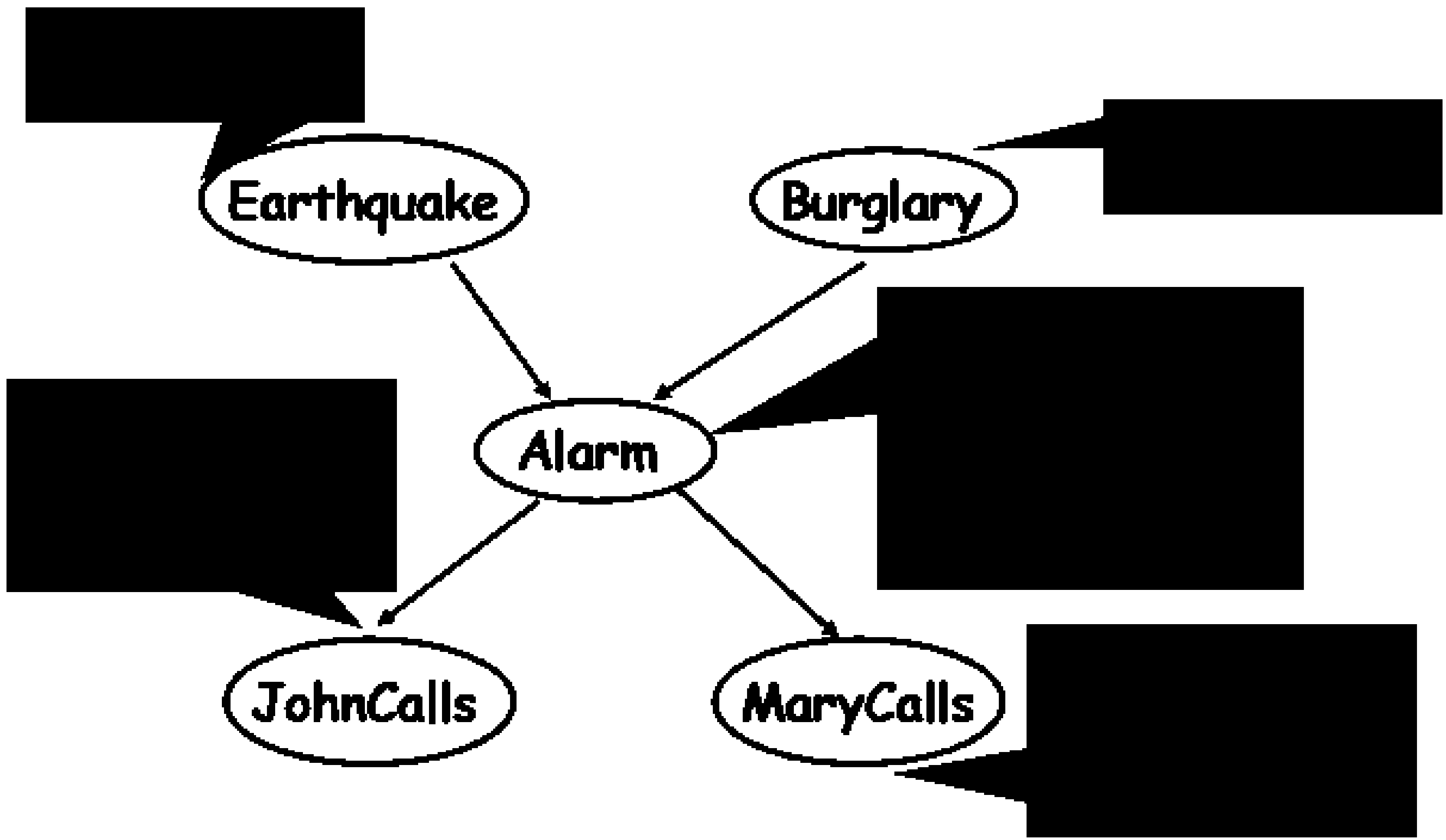
Conditional probability table, or CPT

- Each row in a CPT contains the conditional probability of each node value for a conditioning case.
- A conditioning case is just a possible combination of values for the parent nodes'
- Each row must sum to 1.
- For Boolean variables, once you know that the probability of a true value is p , the probability of false must be 1-p, so we often omit the second number.
- In general, a table for a Boolean variable with k Boolean parents contains 2^k independently specifiable probabilities.
- A node with no parents has only one row, representing the prior probabilities of each possible value of the variable.

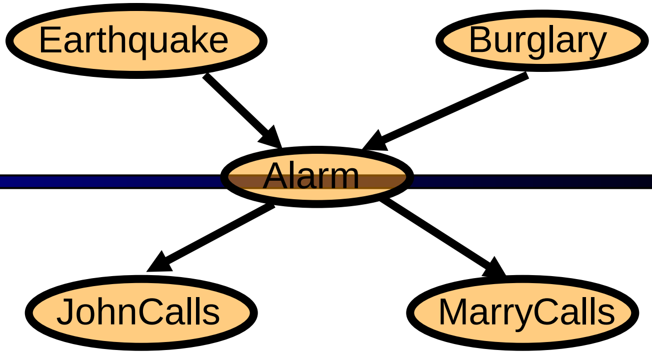
Syntax of BNs

- a set of nodes, one per random variable
- a directed, acyclic graph (link ≈"directly influences")
- a conditional distribution for each node given its parents: P (Xi | Parents (Xi))
  - For discrete variables, conditional probability table (CPT)= distribution over Xi for each combination of parent values

Example Bayes Net: Burglar Alarm



Burglar Alarm Example …



- **If I know if *Alarm*, no other evidence influences my degree of belief in *JohnCalls***
  - $P(J|M,A,E,B) = P(J|A)$
  - also: $P(M|J,A,E,B) = P(M|A)$ and $P(E|B) = P(E)$
- **By the chain rule we have**

$$P(J,M,A,E,B) = P(J|M,A,E,B) \cdot P(M|A,E,B) \cdot P(A|E,B) \cdot P(E|B) \cdot P(B)$$

$$= \mathbf{P(J|A) \cdot P(M|A) \cdot P(A|B,E) \cdot P(E) \cdot P(B)}$$

- **Full joint requires only 10 parameters**

BNs: Qualitative Structure

- Graphical structure of BN reflects conditional independence among variables
- Each variable *X* is a node in the DAG
- Edges denote *direct probabilistic influence*
  - parents of *X* are denoted *Par(X)*
- **Each variable *X* is conditionally independent of all non descendants, given its parents.**
- Graphical test exists for more general independence
  - “Markov Blanket”

Given Parents, X is Independent of Non-Descendants

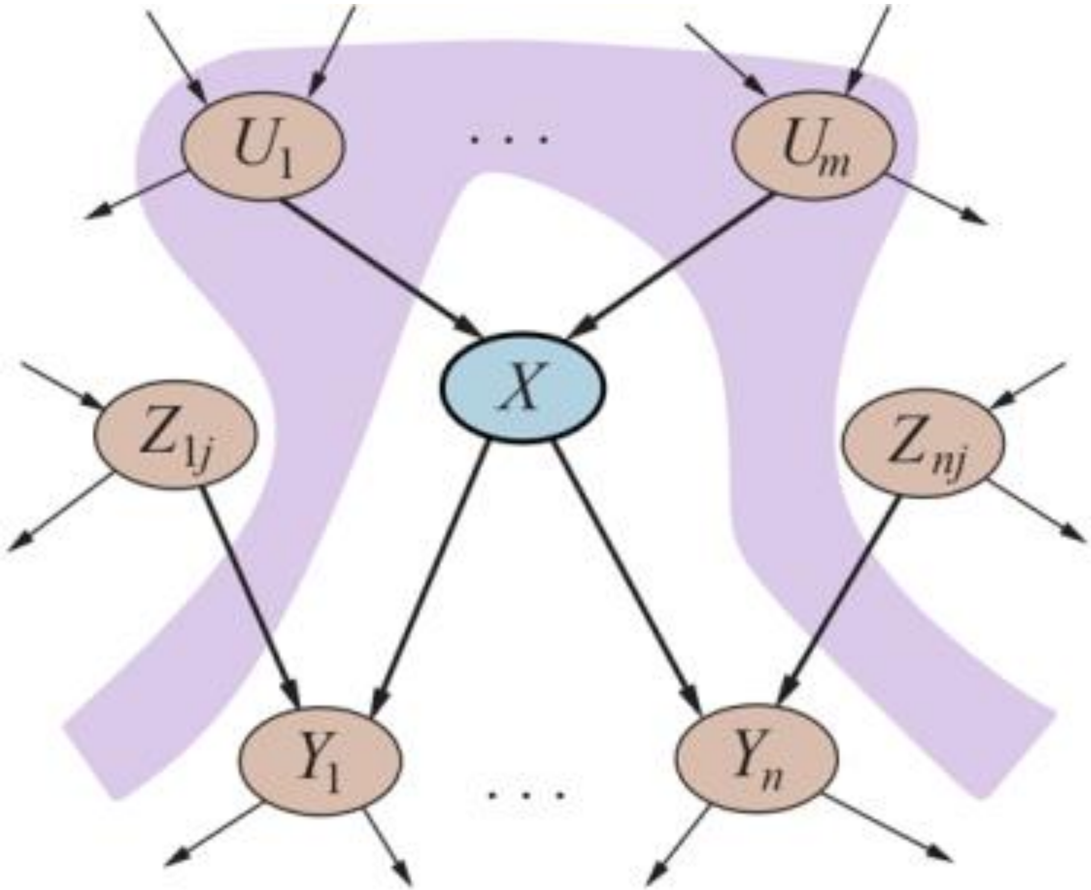
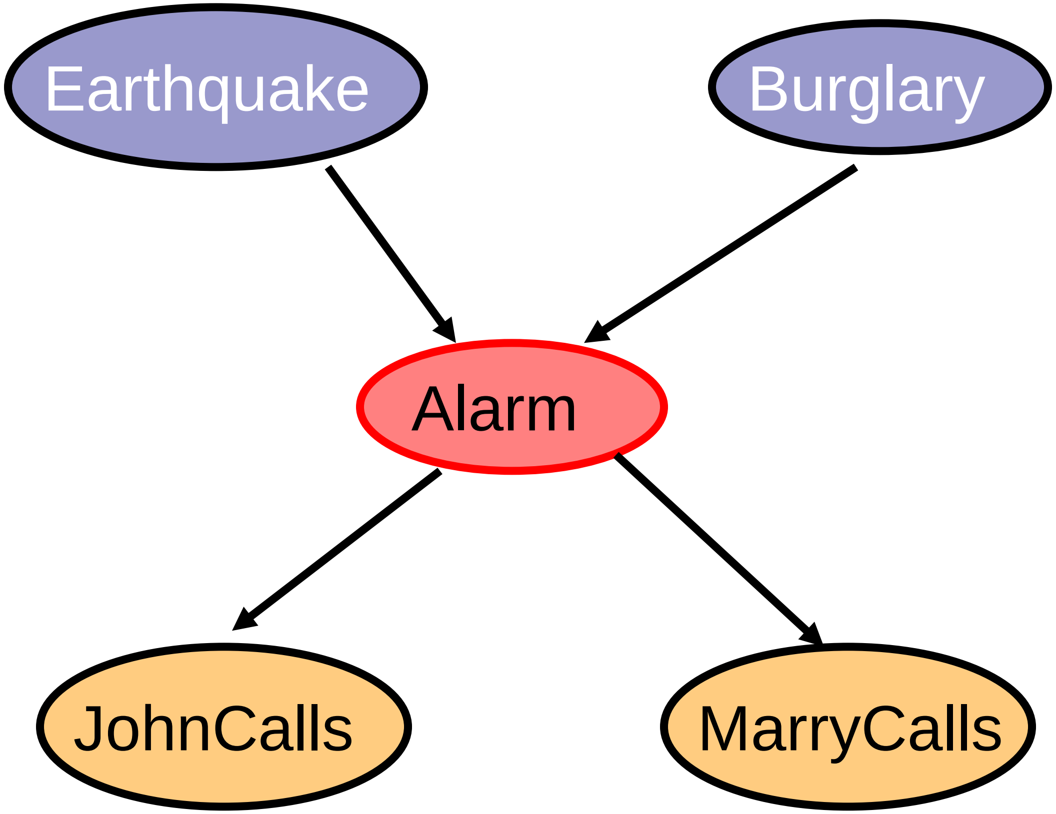
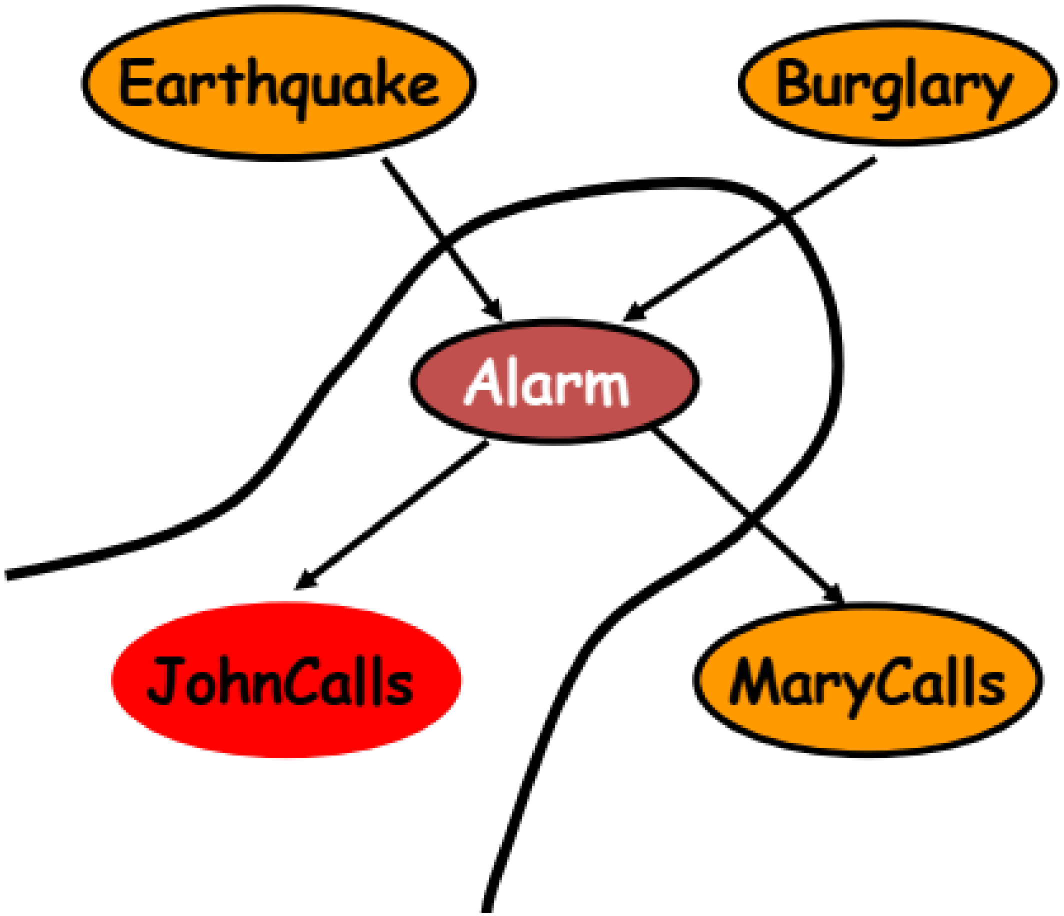


Fig: A node X is conditionally independent of its non-descendants (e.g., the Zij' s) given its parents (the Uis shown in the gray area).

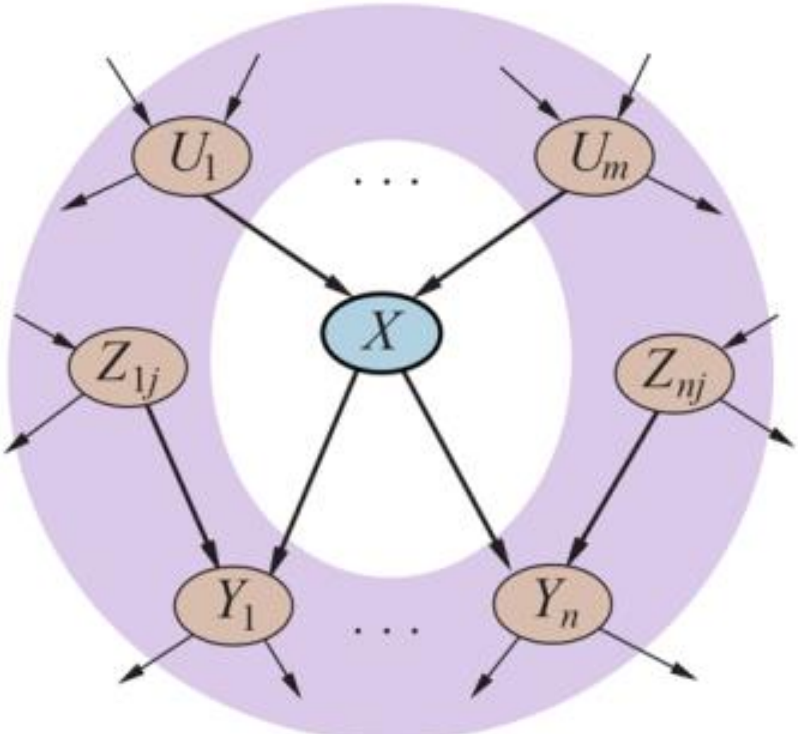
For Example



Example

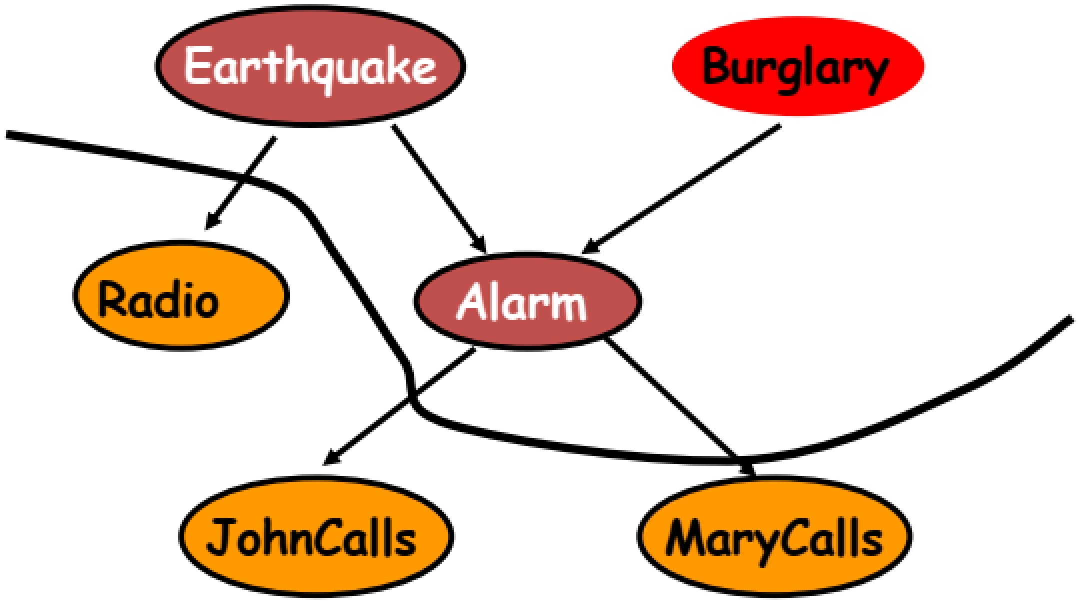


Given Markov Blanket, X is Independent of All Other Nodes

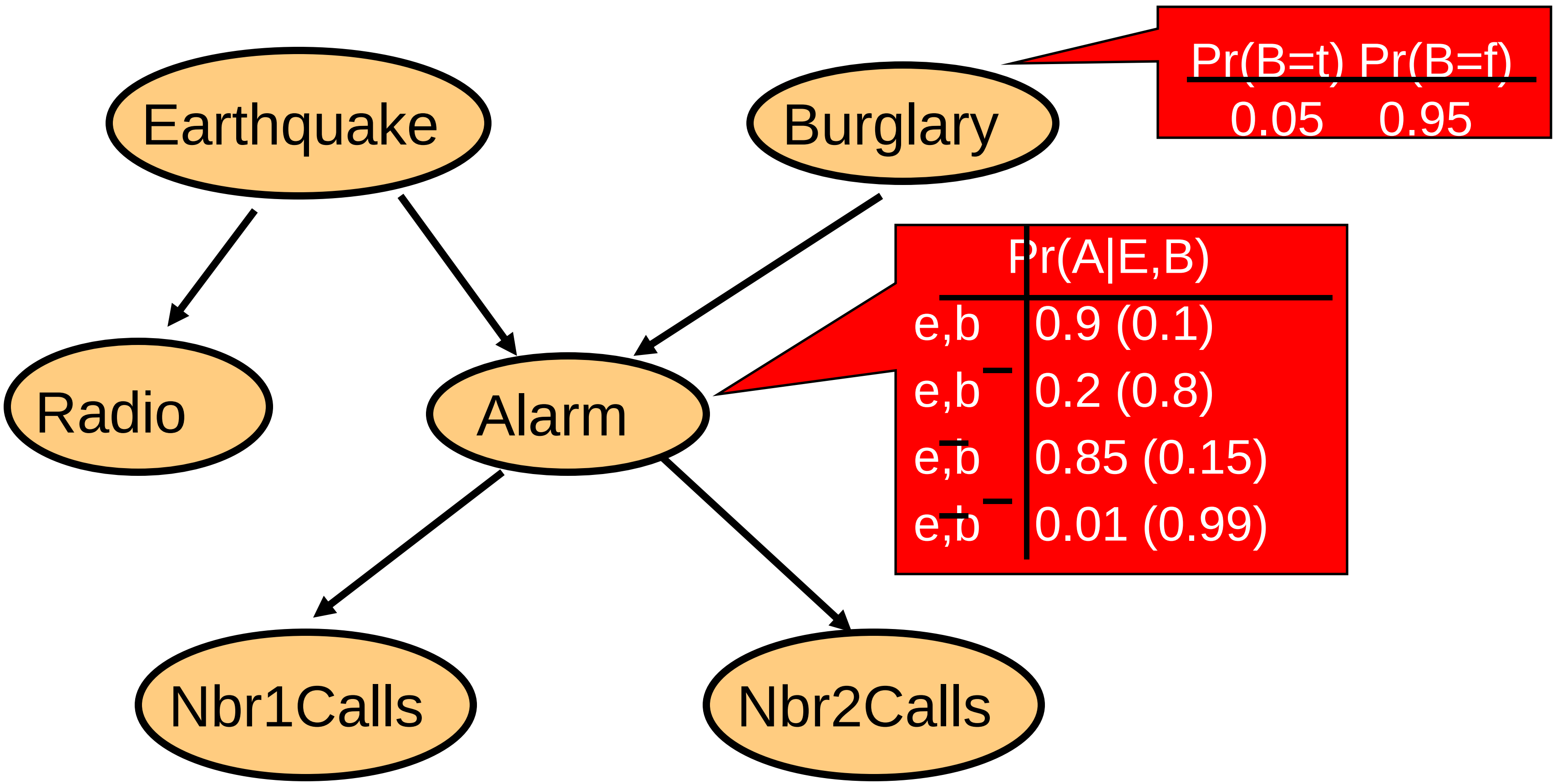


$$MB(X) = Par(X) \cup Childs(X) \cup Par(Childs(X))$$

Example



Conditional Probability Tables



Pr(B=t)	Pr(B=f)
0.05	0.95

	Pr(A|E,B)
e,b	0.9 (0.1)
$e,\bar{b}$	0.2 (0.8)
$\bar{e},b$	0.85 (0.15)
$\bar{e},\bar{b}$	0.01 (0.99)

Conditional Probability Tables

- For complete spec. of joint dist., *quantify* BN

- For each variable X, specify **CPT**: $P(X \mid Par(X))$
  - number of params *locally* exponential in $|Par(X)|$

- If $X_1, X_2, \ldots X_n$ is any topological sort of the network, then we are assured:

$$P(X_n, X_{n-1}, \ldots X_1) = P(X_n \mid X_{n-1}, \ldots X_1) \cdot P(X_{n-1} \mid X_{n-2}, \ldots X_1)$$
$$\ldots \; P(X_2 \mid X_1) \cdot P(X_1)$$

$$= P(X_n \mid Par(X_n)) \cdot P(X_{n-1} \mid Par(X_{n-1})) \ldots P(X_1)$$

Exact Inference in BNs

- The graphical independence representation
  - yields efficient inference schemes
- We generally want to compute
  - *Marginal probability: Pr(Z),* or
  - *Pr(Z|**E**)* where ***E*** is (conjunctive) evidence
    - Z: query variable(s),
    - E: evidence variable(s)
    - everything else: hidden variable
- One simple algorithm:
  - *Inference by enumeration with variable elimination (VE)*

Inference in BNs

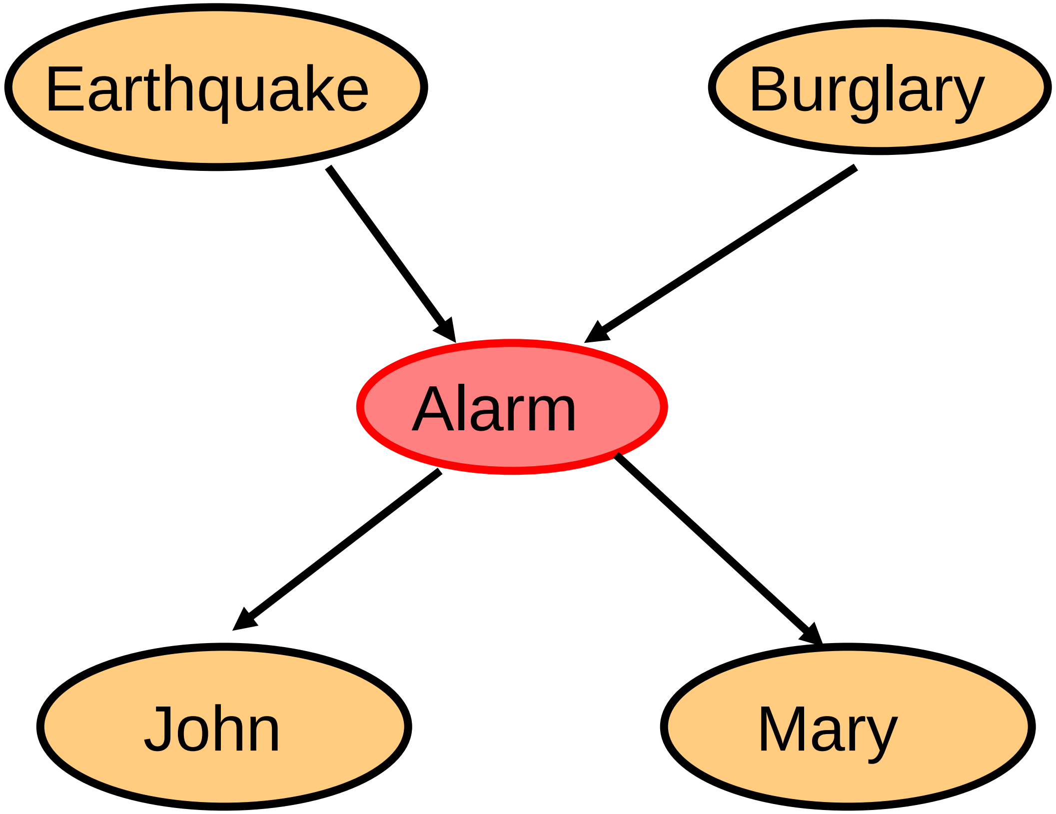
- Let **E** be the list of evidence variables, let **e** be the list of observed values for them, and let y be the remaining unobserved variables (hidden variables). The query P(X | e) can be evaluated as

$$\mathbf{P}(X \mid \mathbf{e}) = \alpha\, \mathbf{P}(X, \mathbf{e}) = \alpha \sum_{\mathbf{y}} \mathbf{P}(X, \mathbf{e}, \mathbf{y})$$

where the summation is over all possible **ys** (i.e., all possible combinations of values of the unobserved variables Y).

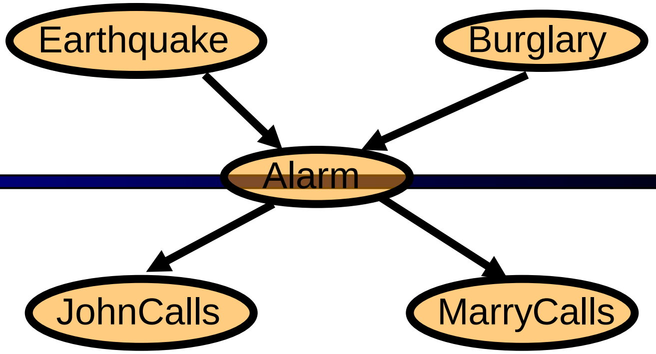
- Now, a Bayes net gives a complete representation of the full joint distribution.
- Therefore, a query can be answered using a Bayes net by computing sums of products of conditional probabilities from the network.

Example: P(B | J=true, M=true)



$$P(B|j,m) = \alpha P(B) \sum_E P(E) \sum_A P(A|B,E)P(j|A)P(m|A)$$

Burglar Alarm Example …



$$P(B|j,m) = \frac{P(B,j,m)}{P(j,m)}$$

$$= \alpha P(B,j,m)$$

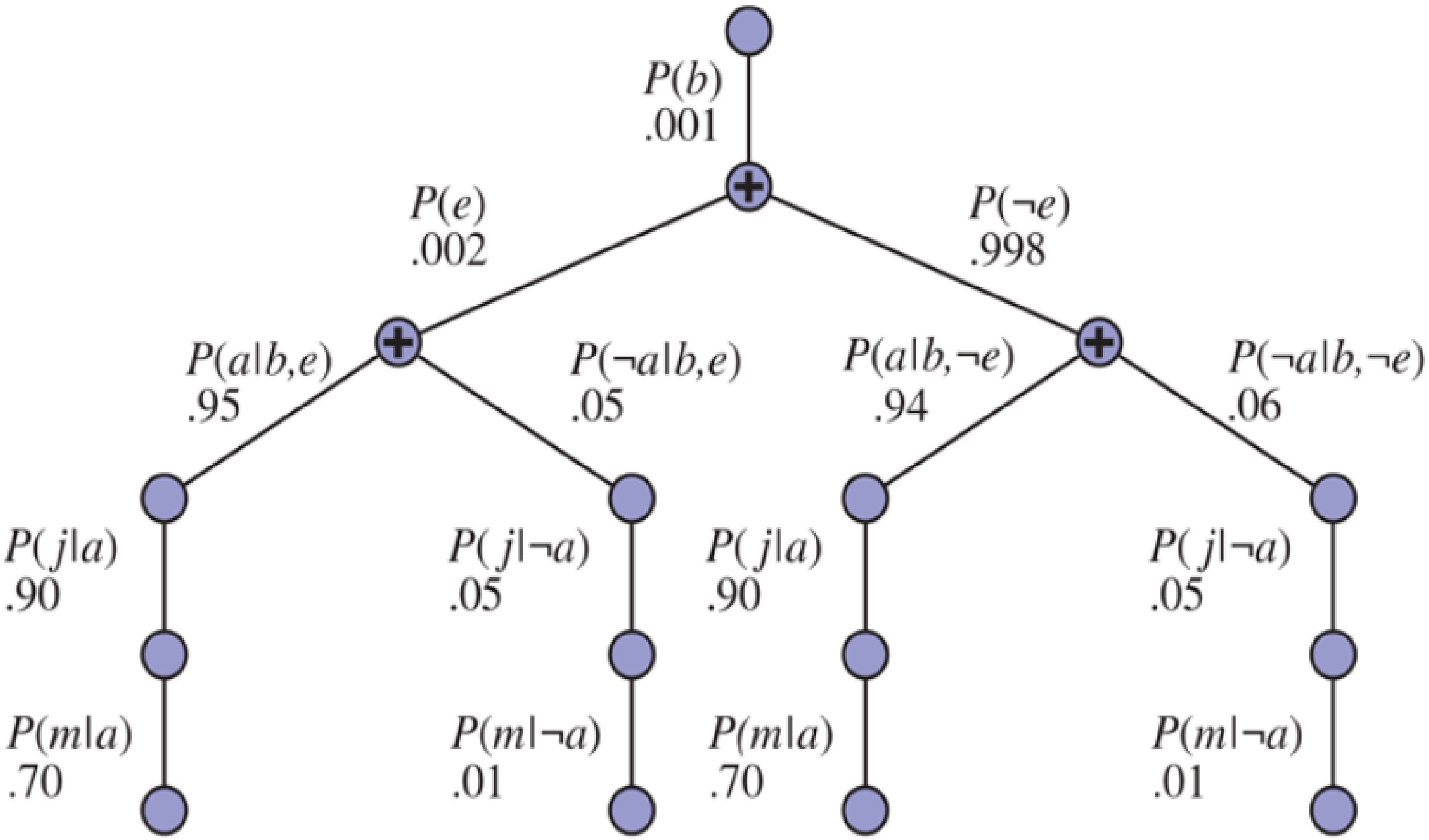
$$= \alpha \textstyle\sum_{E,A} P(B,E,A,j,m)$$

$$= \alpha \sum_{E,A} P(B)P(E)P(A|E,B)P(j|A)P(m|A)$$

$$= \alpha P(B) \sum_{E} P(E) \sum_{A} P(A|E,B)P(j|A)P(m|A)$$

Inference by Enumeration

$P(b|j,m) = \alpha P(b) \sum_E P(E) \sum_A P(A|B,E) P(j|A) P(m|A)$



Dynamic Programming

Variable Elimination

- A *factor* is a function from some set of variables into a specific value: e.g., $f(E,A,B)$
  - CPTs are factors, e.g., $P(A|E,B)$ function of A,E,B
- VE works by *eliminating* all variables in turn until there is a factor with only query variable
- To eliminate a variable:
  - *join* all factors containing that variable (like DB)
  - *sum out* the influence of the variable on new factor

Example of VE: P(J)

$P(J)$

$= \Sigma_{M,A,B,E}\ P(J,M,A,B,E)$

$= \Sigma_{M,A,B,E}\ P(J|A)P(M|A)\ P(B)P(A|B,E)P(E)$

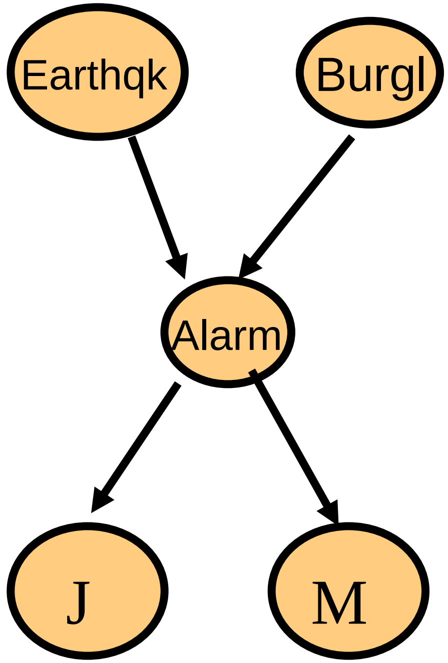
$= \Sigma_{A}P(J|A)\ \Sigma_{M}P(M|A)\ \Sigma_{B}P(B)\ \Sigma_{E}P(A|B,E)P(E)$

$= \Sigma_{A}P(J|A)\ \Sigma_{M}P(M|A)\ \Sigma_{B}P(B)\ f1(A,B)$

$= \Sigma_{A}P(J|A)\ \Sigma_{M}P(M|A)\ f2(A)$

$= \Sigma_{A}P(J|A)\ f3(A)$

$= f4(J)$



Example: P(B | J=true, M=true) using VE

$$\alpha P(B) \sum_{E} P(E) \sum_{A} P(A|E,B) P(j|A) P(m|A)$$

	Pr(J|A)
a	0.9 (0.1)
$\bar{a}$	0.05 (0.95)

	Pr(M|A)
a	0.7 (0.3)
$\bar{a}$	0.01 (0.99)

	Pr(j|A)P(m|A)
a	0.9x0.7
$\bar{a}$	0.05x0.01

$$\alpha P(B) \sum_E P(E) \sum_A P(A|E,B) f1(A)$$

	Pr(J|A)
a	0.9 (0.1)
$\bar{a}$	0.05 (0.95)

	Pr(M|A)
a	0.7 (0.3)
$\bar{a}$	0.01 (0.99)

	f1(A)
a	0.63
$\bar{a}$	0.0005

$$\alpha P(B) \sum_{E} P(E) \sum_{A} P(A|E,B) f1(A)$$

	f1(A)
a	0.63
$\bar{a}$	0.0005

	Pr(A|E,B)
e,b	0.95 (0.05)
e,$\bar{b}$	0.29 (0.71)
$\bar{e}$,b	0.94 (0.06)
$\bar{e}$,$\bar{b}$	0.001 (0.999)

e,b	0.95x0.63+0.05x0.0005
e,$\bar{b}$	0.29x0.63+0.71x0.0005
$\bar{e}$,b	0.94x0.63+0.06x0.0005
$\bar{e}$,$\bar{b}$	0.001x0.63+0.999x0.0005

$$\alpha P(B) \sum_{E} P(E) f2(E,B)$$

	f1(A)
a	0.63
$\bar{a}$	0.0005

	Pr(A|E,B)
e,b	0.95 (0.05)
e,$\bar{b}$	0.29 (0.71)
$\bar{e}$,b	0.94 (0.06)
$\bar{e}$,$\bar{b}$	0.001 (0.999)

	f2(E,B)
e,b	0.60
e,$\bar{b}$	0.18
$\bar{e}$,b	0.59
$\bar{e}$,$\bar{b}$	0.001

$$\alpha P(B) \sum_{E} P(E) f2(E,B)$$

Pr(E=t)	Pr(E=f)
0.002	0.998

Pr(B=t)	Pr(B=f)
0.001	0.999

	f2(E,B)
e,b	0.60
$e,\bar{b}$	0.18
$\bar{e},b$	0.59
$\bar{e},\bar{b}$	0.001

b	0.60x0.002x0.001 + 0.59x0.998x0.001
$\bar{b}$	0.18x0.002x0.999 + 0.001x0.998x0.999

αf3(B)

Pr(E=t)	Pr(E=f)
0.002	0.998

Pr(B=t)	Pr(B=f)
0.001	0.999

	f2(E,B)
e,b	0.60
$e,\bar{b}$	0.18
$\bar{e},b$	0.59
$\bar{e},\bar{b}$	0.001

	f3(B)
b	0.0006
$\bar{b}$	0.0013

$$\alpha f3(B) \rightarrow P(B|j,m)$$

	f3(B)
b	0.0006
$\bar{b}$	0.0013

$$N = 0.0006+0.0013$$
$$= 0.0019$$

$$\alpha f3(B) \rightarrow P(B|j, m)$$

	f3(B)
b	0.0006
$\bar{b}$	0.0013

$$N = 0.0006+0.0013$$
$$= 0.0019$$

	P(B|j,m)
b	0.32
$\bar{b}$	0.68

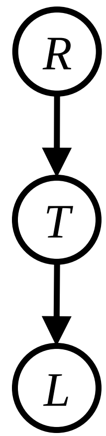
Example: Traffic Domain

- Random Variables
  - R: Raining
  - T: Traffic
  - L: Late for class!

$$P(L) = ?$$

$$= \sum_{r,t} P(r, t, L)$$

$$= \sum_{r,t} P(r)P(t|r)P(L|t)$$



$P(R)$

+r	0.1
-r	0.9

$P(T|R)$

+r	+t	0.8
+r	-t	0.2
-r	+t	0.1
-r	-t	0.9

$P(L|T)$

+t	+l	0.3
+t	-l	0.7
-t	+l	0.1
-t	-l	0.9